GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG) ENZYMES, THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE THEREWITH

Cross Reference to Related Application

This Application claims the benefits of US Provisional Application number 60/083,768, filed May 1, 1998. The entire disclose of US Provisional Application 60/083,768 is incorporated herein by reference.

Statement Regarding Federally Sponsored Research Or Development

The present invention was supported in part by the National Institutes of Health (Grant CA43894). The United States Government may have certain rights in the invention.

Technical Field

The present invention relates to poly(ADP-ribose) glycohydrolases (PARGs) and peptides having poly(ADP-ribose) glycohydrolase activity. In addition, the invention also relates to antibodies, including monoclonal antibodies and antibody fragments, that have specific interaction with epitopes present on poly(ADP-ribose) glycohydrolases. Methods of treatment and diagnosis using the poly(ADP-ribose) glycohydrolases, and antibodies specific for poly(ADP-ribose) glycohydrolases are disclosed. The present invention has implications for the treatment of neoplastic disorder, reperfusion following ischemia, neurological disorders, and related conditions.

Background of the Invention

Genomic damage, if left unrepaired, can lead to malignant transformation, or cell death by senescence (aging), necrosis or apoptosis. Among the variables that can affect the ultimate biological consequence of DNA damage to a particular cell are (i) the amount, type, and location of the DNA damage and (ii) the efficiency and bioavailability of the cellular DNA repair mechanism.

The activation of poly(ADP-ribose) polymerase (PARP) by DNA strand breaks is often one of the first cellular responses to DNA damage. PARP catalyzes the conversion of nicotinamide adenine dinucleotide (NAD) to multi-branched polymers containing up to 200 ADP-ribose residues. Increases in polymer levels of more than 100-fold may occur within minutes of DNA damage. Once synthesized, polymers are rapidly turned over, being converted to free ADP-ribose by the action of poly(ADP-ribose) glycohydrolase (PARG) (*1*). An ADP-ribosyl protein lyase has been proposed to catalyze removal of protein-proximal ADP-ribose monomers (*2*). Figure 1 illustrates these processes schematically.

The process of activating PARP upon DNA damage can rapidly lead to energy depletion because each ADP-ribose unit transferred by PARP consumes one molecule of NAD, which in turn, requires six molecules of ATP to regenerate NAD. Additionally, NAD is a key carrier of electrons needed to generate ATP via electron transport and oxidative phosphorylation or by glycolysis. The overactivation of PARP due to substantial DNA damage can significantly deplete the cellular pools of NAD and ATP (*3*). ADP-ribose polymer metabolism, and thus PARP and PARG have been linked to the enhancement of DNA repair (*4*), limitation of malignant transformation (*5*), enhancement of necrotic cell death (*6*), and involvement in programmed cell death (*7*). To date, studies of the structure and function of the enzymes of ADP-ribose polymer metabolism have been mainly limited to PARP (*8*). Little is known about the function and regulation of PARG.

Brief Summary of the Invention

As embodied and broadly described herein, the present invention is directed to nucleic acids molecules, peptides, methods, vectors and antibodies that are related to the poly(ADP-ribose) glycohydrolase (PARG) enzyme.

One embodiment of the invention is directed to an isolated and purified nucleic acid molecule or nucleic acid molecule analog comprising a sequence that encodes a polypeptide having poly(ADP-ribose) glycohydrolase (PARG) activity. The nucleic acid molecule may encode the complete full-length PARG gene or a fragment of the PARG gene. The nucleic acid molecule may be DNA, RNA or peptide nucleic acid (PNA). The nucleic acid molecule can be linear, such as, for example, an isolated fragment or a linear phage DNA. In addition, the isolated nucleic acid molecule may be circular, such as for example in a plasmid. The nucleic

acid molecule may also be a single stranded DNA or RNA such as the single stranded DNA or RNA in a single stranded DNA virus or single stranded RNA virus. The nucleic acid molecule may be of yeast, insect or mammalian origin.

The nucleic acid molecule of the invention, may be of mammalian origin, such as, for example of bovine or murine origin. In a preferred embodiment of the invention, the nucleic acid molecule may be of human origin. While the sequence of the nucleic acid molecule is of mammalian origin, the nucleic acid molecule may be replicated in another organism such as an insert in a viral genome, a plasmid in a bacterium or a 2-micron plasmid in a yeast.

Preferably, the nucleic acid molecule has, a high degree of sequence similarity with a sequence shown in SEQ ID NO: 1 (Genbank Accession Number U78975), SEQ ID NO: 3 (Genbank Accession Number AF005043), SEQ ID NO: 5 (Genbank Accession Number AF079557), SEQ ID NO: 7 (Genbank Accession Number AF079556) or SEQ ID NO: 9 (Genbank Accession Number CEF20C5). The high degree of sequence similarity may be, for example, about 70%, preferably about 80%, even more preferably about 90% and most preferably substantially identical such as for example about 100% identity.

The nucleic acid molecule that encodes a polypeptide having poly(ADP-ribose) glycohydrolase (PARG) activity may be single or double stranded nucleic acid molecule of any length such as, for example, about 20 bases in length, about 30 bases in length, about 40 bases in length, about 50 bases in length, about 100 bases in length, about 200 bases in length, about 500 bases in length, about 1000 bases in length, about 1500 bases in length, about 2000 bases in length, about 3000 bases in length. It is understood that "bases" in this patent application means "basepairs" when referring to double stranded nucleic acid molecules and bases when referring to single stranded nucleic acid molecules. In a preferred embodiment of the invention, the nucleic acid molecule may be at least about 1000 base or basepairs long and have at least about 80% sequence similarity with a sequence shown in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7 or SEQ ID NO: 9.

In one embodiment of the invention, the nucleic acid molecule may have sequence similarity to one region of the PARG sequence. The region may be, for example, from about base residue 2113 to about residue 3105 of SEQ ID NO: 3. Alternatively, the region may be, from residue 1240 to about residue 3105 of SEQ ID NO: 3 or from residue 175 to about residue 3105 of SEQ ID NO: 3.

Another embodiment of the invention is directed to the expression and overexpression of PARG in a cell. Expression vectors may mediate the expression of a polypeptide with poly (ADP-ribose) glycohydrolase (PARG) enzyme activity. Expression systems and expression vectors are known in the art. For example, one expression vector may comprise a regulatory sequence which is operatively linked to a nucleotide sequence at least about 1000 base pairs in length, which has at least 70% sequence similarity with a sequence shown in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7 or SEQ ID NO: 9. In a preferred embodiment, the sequence similarity is at least about 80% identity, more preferably at least about 90% identity and most preferably about 100% identity. The expression vector may be any expression vector that is capable of directing expression of a gene in a host cell including, prokaryotic, eukaryotic, mammalian and viral vector. Examples of such vectors include pCMV-Script cytomeglovirus expression vectors for expression in mammalian cells, pESP and pESC vectors for expression in *S. pombe* and *S. cerevesiae*, pET vectors for expression in bacteria, pSPUTK vectors for high-level transient expression, and pPbac and pMbac vectors for expression in fall army worm (SF9) cells. Such vectors are available commercially from suppliers such as, for example, Invitrogen (Carlsbad, CA) or Stratagene (La Jolla, CA). In the use of viral vectors, it is understood that defective viral vectors -- vectors that are genetically engineered to deliver a gene or gene product to a host but which cannot replicate in a host is preferred. Procedures for the practice of *in vitro* and *in vivo* expression are well known to those of skill in the art and are further available with the specific expression products and cell lines from commercial suppliers.

Another embodiment of the invention is directed to a host cell transformed with a vector containing a nucleic acid molecule with a sequence that encodes a polypeptide having poly(ADP-ribose) glycohydrolase (PARG) activity. The host cell may be any eukaryotic or prokaryotic cell such as, for example a human, murine, rattus, bovine, insect, yeast or bacteria. Specific cell lines are well known to those of skill in the art and are available from suppliers such as the American Tissue Type Collection (ATCC, Manassas, VA) and Stratagene (La Jolla, CA) and the like. A preferred embodiment of the invention is directed to cells transformed with the PARG expression vector which shows an elevated level of PARG relative to non-transformed cells. Especially preferred are cells transformed with an inducible PARG expression vector that have normal or slightly elevated PARG levels before induction and have significantly elevated PARG levels after induction.

An embodiment of the invention is directed to an isolated protein having poly(ADP-ribose) glycohydrolase (PARG) activity. The protein may comprise an amino acid sequence with at least 70% sequence similarity with a sequence shown in SEQ ID NO: 2 (Genbank Accession Number U78975), SEQ ID NO: 4 (Genbank Accession Number AF005043), SEQ ID NO: 6 (Genbank Accession Number AF079557), SEQ ID NO: 8 (Genbank Accession Number AF079556), or SEQ ID NO: 10 (Genbank Accession Number CEF20C5). The sequence similarity is preferably at least about 80%, more preferably at least about 90% and most preferably substantially identical with a sequence shown in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, or SEQ ID NO: 10. In a preferred embodiment of the invention, the preferred isolated protein having poly(ADP-ribose) glycohydrolase (PARG) activity and has a molecular weight greater than about 100 kDa.

Another embodiment of the invention is directed to an oligonucleotide which is greater than about 10 bases in length and less than about 1000 bases in length which is complementary to a sequence shown SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, or SEQ ID NO: 9. The oligonucleotide may be, for example, greater than about 20 bases in length, greater than about 30 bases in length, greater than about 40 bases in length, greater than about 50 bases in length, greater than about 100 bases in length, greater than about 200 bases in length or greater than about 300 bases in length. The oligonucleotide, which may be optionally labeled with a detectable marker, may be DNA, RNA or PNA. A detectable marker may be, for example, a radioactive isotope such as ^{32}P or ^{125}I, an epitope such as FLAG.

One preferred oligonucleotide is an antisense oligonucleotide directed to the mRNA of PARG. Antisense oligonucleotide as a method of suppression is well known to those in the art. For example, the phosphorothioate oligonucleotide, ISIS 2922, has been shown to be effective against cytomeglovirus retinitis in AIDS patients (*9*). It is thus well known that oligonucleotides, when administered to animals and humans, can have a useful therapeutic effect. In a preferred embodiment, the oligonucleotide is at least about 10 nucleotides in length, such as, greater than about 20 bases in length, greater than about 30 bases in length, greater than about 40 bases in length, greater than about 50 bases in length, greater than about 100 bases in length, greater than about 200 bases in length or greater than about 300 bases in length. In another preferred embodiment, the oligonucleotide has a ribozyme activity.

Another embodiment of the invention is directed to an isolated polypeptide of at least 6 amino acid residues in length and having a molecular weight less than about 65 kDa, which has at least about 80% sequence similarity with a sequence shown in any one of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 or SEQ ID NO: 10. The polypeptide may be, for example, at least about 10 amino acids in length, at least about 20 amino acids in length, at least about 30 amino acids in length, at least about 40 amino acids in length, at least about 50 amino acids in length, at least about 75 amino acids in length, at least about 100 amino acids in length, at least about 150 amino acids in length, at least about 250 amino acids in length or at least about 500 amino acids in length or more.

In a preferred embodiment, the polypeptide has a molecular weight less than about 40 kDa and has at least about 90% sequence similarity with a sequence shown in any one of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 or SEQ ID NO: 10. The polypeptide preferably has poly(ADP-ribose) glycohydrolase (PARG) activity or is immunogenic and elicits antibodies immunoreactive with a poly(ADP-ribose) glycohydrolase (PARG) enzyme. In a more preferred embodiment, the polypeptide comprises an amino acid sequence substantially identical with SEQ ID NO: 4 from about residue 647 to about residue 977.

Another embodiment of the invention is directed to an isolated polypeptide of at least 10 amino acid residues in length and which has at least about 80% sequence similarity with a sequence shown in any one of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 or SEQ ID NO: 10. Preferably, the polypeptide is at least about 20 amino acids in length, such as, for example at least about 30 amino acids, about 40 amino acids, about 50 amino acids, about 100 amino acids, about 200 amino acids and about 500 amino acids in length.

Another embodiment of the invention is directed to an antibody immunoreactive with an isolated polypeptide of at least about 6 amino acid residues in length and having a molecular weight less than about 65 kDa, which has at least about 80% sequence similarity with a sequence shown in any one of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 or SEQ ID NO: 10. In a preferred embodiment, antibody is immunoreactive with a polypeptide with a molecular weight less than about 40 kDa and has at least about 90% sequence similarity with a sequence shown in any one of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 or SEQ ID NO: 10. In another preferred embodiment, the antibody is immunoreactive with a

polypeptide comprising an amino acid sequence substantially identical with SEQ ID NO: 4 from about residue 647 to about residue 977.

Another embodiment of the invention is directed to a method of detecting a polypeptide having PARG activity comprising the steps of contacting the polypeptide with an antibody immunoreactive with an isolated polypeptide of at least about 6 amino acid residues in length and having a molecular weight less than about 65 kDa, which has at least about 80% sequence similarity with a sequence shown in any one of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 or SEQ ID NO: 10, and determining whether the antibody immunoreacts with the polypeptide.

Another embodiment of the invention is directed to a method of preventing, treating, or ameliorating a disease condition or disorder in an individual comprising the step of administering a therapeutically effective amount of a poly(ADP-ribose) glycohydrolase (PARG) inhibitor or activator to the individual. The disease condition or disorder may be any condition associated with responses to DNA damage, examples of which include a neoplastic disorder, a myocardial infarction, a vascular stroke or a neurodegenerative disorder. The PARG inhibitor or activator may be a small molecule inhibitor or activator of PARG or may be an antisense oligonucleotide that can hybridize *in vivo* to messenger RNA encoded by a PARG gene. PARG based treatment may be directed to new methods for preventing, treating or ameliorating disorders associated with DNA damage. These disorders include neoplastic disorders, inborn genetic errors, myocardial infarctions, vascular strokes, aging, and neurodegenerative disorders such as Alzheimer's disease, Huntington's disease, Parkinson's disease, and neurotoxicity generally.

Another embodiment of the invention is directed to the identification of novel PARG modulators which can activate or inhibit DNA repair and/or apoptosis. A PARG modulator is a compound that can activate or inhibit PARG. These modulators are preferably more efficacious and do not have the known side effects of present modulators. One method of identifying an agent that inhibits or activates poly(ADP-ribose) glycohydrolase (PARG) activity comprise the steps of providing a liquid medium that contains a polypeptide having PARG activity contacting the polypeptide with a candidate agent, in the presence of a reference compound having affinity for the polypeptide, under predetermined assay conditions, and determining the affinity of the candidate agent for the polypeptide relative to the reference compound. Thus, the modulation activity of the candidate agent relative to the reference

compound is determined. In this method, the polypeptide may be immobilized on a solid support. Further, the polypeptide may be generated *in vitro* by culturing a cell transformed with a nucleic acid molecule encoding PARG under conditions effective to express the polypeptide.

Another embodiment of the invention is directed to a method of identifying a mutant PARG allele in an individual comprising the step of obtaining genomic material from the individual; digesting the genomic material with a restriction enzyme having a recognition site inclusive of the mutant allele; fractionating the restriction fragments obtained from the digestion; and comparing the fractionation pattern with that obtained for a normal allele, thereby determining the presence or absence of the mutant allele. The fractionating step may be performed with electrophoresis.

Another embodiment of the invention is directed to a method of identifying a mutant PARG allele in an individual comprising the steps of hybridizing an oligonucleotide with genomic material from the individual, which oligonucleotide hybridizes under predetermined hybridization conditions to a region immediately 5' of a predetermined mutation site in the PARG alleles with the 3' terminus of the oligonucleotide complementary to an unmutated PARG allele; extending the oligonucleotide using PCR amplification; and determining the degree to which extension occurs, thereby determining the presence or absence of the mutant allele. The PCR extension reaction may be performed at a temperature above about 50°C. The determination may be performed by conducting electrophoresis (using for example, acrylamide at about 4% to about 10% or agarose and low melting temperature agarose from about 0.8% to about 4%) on the products of PCR amplification.

Another embodiment of the invention is directed to a method of screening molecules for PARG modulating activity (inhibition or activation) comprising the steps of providing a purified PARG enzyme; assaying the enzyme in the presence of a molecule to be screened; and comparing the activity of the PARG enzyme in the presence of the molecule to the activity of the PARG enzyme in the absence of the molecule.

Another embodiment of the invention is directed to a method of gene therapy comprising the step of delivering an oligonucleotide having a sequence complementary to at least a portion of a polynucleotide encoding a PARG enzyme to a cell to be treated. In the method, the oligonucleotide may have a sequence complementary to a sequence encoding a C-terminal

portion of a PARG enzyme. Further, in the gene therapy method, the oligonucleotide may further comprise a ribozyme.

Another embodiment of the invention is directed to a method of delivering to a cell surface, an oligonucleotide having a sequence complementary to at least a portion of a polynucleotide encoding a PARG enzyme to a cell to be treated. In the method, the oligonucleotide may have a sequence complementary to a sequence encoding a C-terminal portion of a PARG enzyme. Further, in the method, the oligonucleotide may further comprise a ribozyme. The portion of a polynucleotide encoding a PARG enzyme may be, for example, the polynucleotide encoding the N terminus third of PARG, the middle third of PARG, or the C terminus third of PARG. The portion of a polynucleotide may encode a smaller part of PARG such as the N terminus 10% of PARG, the C terminus 10% of PARG, or any 10% portion in between such as from 10% to 20%, from 20% to 30%, from 30% to 40%, from 40% to 50%, from 50% to 60%, from 60% to 70%, from 70% to 80%, from 80% to 90%. The percent value used means a percent of the linear amino acid sequence. Thus, for a 1000 amino acid protein, the N terminus 10 percent is from amino acid 1 to 100; 10% to 20% percent would be from amino acid 100 to 200 and so on. For a 970 amino acid protein, the N terminal 10% would be from amino acid 1 to 97; 10% to 20% would be from amino acids 98 to 194 amino acids.

Another embodiment of the invention is directed to a method of sensitizing a cell to a chemotherapeutic agent comprising the step of contacting the cell with a molecule that modulates the activity of a PARG enzyme. The molecule may be an oligonucleotide having a sequence complementary to at least a portion of a polynucleotide encoding a PARG enzyme. For example, the oligonucleotide may have a sequence complementary to a sequence encoding a C-terminal portion of a PARG enzyme. The portion of a polynucleotide encoding a PARG enzyme may be, for example, the polynucleotide encoding the N terminus third of PARG, the middle third of PARG, or the C terminus third of PARG. The portion of a polynucleotide may encode a smaller part of PARG such as the N terminus 10% of PARG, the C terminus 10% of PARG, or any 10% portion in between such as from 10% to 20%, from 20% to 30%, from 30% to 40%, from 40% to 50%, from 50% to 60%, from 60% to 70%, from 70% to 80%, from 80% to 90%. The oligonucleotide may further comprise a ribozyme. The method may be used, for example, as a method of treating a diseased cell characterized by the presence of DNA strand breaks. In the

treatment, the cell is contacted with a molecule that modulates an enzymatic activity of a PARG enzyme.

Another embodiment of the invention is directed to a pharmaceutical composition comprising an oligonucleotide having a sequence complementary to at least a portion of a polynucleotide encoding a PARG enzyme. The produced molecule may be an oligonucleotide having a sequence complementary to at least a portion of a polynucleotide encoding a PARG enzyme. For example, the oligonucleotide may have a sequence complementary to a sequence encoding a C-terminal portion of a PARG enzyme. The oligonucleotide may comprise a ribozyme activity.

Another embodiment of the invention is directed to a virus that causes the production of an oligonucleotide having a sequence complementary to a polynucleotide encoding a PARG enzyme. This may be, for example, a viral vector which after the infection of a host cell, causes the production of an antisense RNA of PARG. The molecule may be an oligonucleotide having a sequence complementary to at least a portion of a polynucleotide encoding a PARG enzyme. For example, the oligonucleotide may have a sequence complementary to a sequence encoding a C-terminal portion of a PARG enzyme. The oligonucleotide may further comprise a ribozyme activity.

Other embodiments and advantages of the invention are set forth, in part, in the description that follows and, in part, will be obvious from this description and may be learned from the practice of the invention.

Brief Description of the Drawings

Figure 1 depicts the cellular biochemical process that occurs after DNA damage.

Figure 2 depicts the SDS-PAGE analysis of purified bovine thymus PARG.

Figure 3 depicts the alignment of the DNA sequences of two PCR products and eight λgt11 cDNA clones used to identify the cDNA coding for bovine PARG.

Figure 4 depicts a northern blot analysis of bovine kidney cells mRNA transcripts.

Figure 5 depicts an alignment of the putative bipartite NLS of bovine, human, and murine PARG and comparison with the bipartite NLS of PARP from different organisms.

Figure 6 depicts expression of bPARG enzyme activity in *E. coli* (*10*).

Figure 7 depicts a Southern blot analysis of bovine DNA probed with PARG cDNA.

Figure 8 depicts activity gel autoradiogram of *E. coli* expressed bovine PARG.

Figure 9 depicts the analysis by anion exchange HPLC of material released from ADP-ribose polymers by PARG action.

Figure 10 depicts the SDS-PAGE analysis of the purification of *E. coli* expressed GST-PARG.

Figure 11 depicts a schematic representation of the portions of the bovine PARG cDNA expressed as GST fusion constructs.

Figure 12 depicts the cloning of the 1.8 kb PCR *Eco*RI fragment encoding for the 65 kDa catalytic domain of PARG.

Figure 13 depicts an autoradiogram of an activity gel of GST-PARG fusion constructs expressed in *E. coli* and PARG expressed in baculovirus.

Figure 14 depicts a schematic representation of the strategy used to isolate cDNA molecules encoding PARG from various organisms.

Figure 15 depicts the domain organization of PARGs from different organisms.

Figure 16 depicts an amino acid sequence alignment of bovine, murine, human, drosophila and C. elegans PARG enzymes.

Figure 17 depicts a western blot of recombinant PARGs.

Figure 18 depicts western blots of natural and recombinant expressed PARG.

Figure 19 depicts the characterization of PARG by Western Blot in mouse cells of different PARP genotypes.

Figure 20 depicts a partial restriction map of the mouse PARG locus.

Figure 21 depicts a schematic representation of the strategy used to create PARG knockout mice.

Detailed Description of the Invention

DEFINITIONS

List of Abbreviations:

ADP	adenosine diphosphate
ADPR	ADP-ribose
AMP	adenosine monophosphate
ASPCR	allele-specific PCR
bp	base pair(s)
*b*PARG	bovine PARG
CePARG	*C. elegans* PARG
dPARG	*Drosophila melanogaster* PARG
DTT	dithiothreitol
GSH-Sepharose	Glutathione-Sepharose 4B
GST	glutathione-S transferase
hPARG	human PARG
HPLC	high pressure liquid chromatography
ICE	interleukin-1 b converting enzyme
IPTG	isopropyl-ß-D-thiogalactoside
kb	kilobase pair(s)
MDBK	Madin-Darby bovine kidney cells
*m*PARG	murine PARG
NAD	nicotinamide adenine dinucloetide
NLS	nuclear location signal
PADPR DHB-Sepharose	poly(ADP-ribose)-dihydroxyboronyl-Sepharose
PAGE	polyacrylamide-gel electrophoresis
PARG	poly(ADP-ribose) glycohydrolase
PARP	poly(ADP-ribose) polymerase [EC 2.4.2.30]
PCR	polymerase chain reaction
PEG-6,000	polyethylene glycol 6,000
PEG	polyethylene glycol
PMSF	phenylmethylsulfonyl fluoride
PR-AMP	phosphoribosyl-adenosine monophosphate
RFLP	restriction fragment length polymorphism
SDS	sodium dodecyl sulfate
SSCP	single-strand conformation polymorphism
TPCK	Trypsin: L-1-tosylamido-2-phenylethyl chloromethyl ketone.

An "agonist" as defined herein refers to a molecule which, when bound to PARG, increases or prolongs the effect of PARG. Agonist may include proteins, nucleic acid molecules, carbohydrates, or any other molecules that bind to and modulate the effect of PARG.

An "allele" or "allelic sequence", as defined herein refers to an alternative form of PARG. Alleles may result from at least one mutation in the nucleic acid molecule sequence and may result in altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to alleles, are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

An "ortholog" as defined herein refers to a nucleotide or amino acid sequence that is related to a reference nucleotide or amino acid sequence through speciation, and is therefore identical or structurally similar to the reference sequence.

A given nucleotide or amino acid sequence is said to be "substantially identical" with another sequence when the compared sequences have the same residues in the same order, excepting for any degeneracy (nucleotides) and conservative substitutions (amino acids).

A "regulatory sequence" of an expression vector is a DNA sequence necessary for inducing transcription of a gene, and includes a functional promoter and/or enhancer sequence. The term "operatively linked" as used herein means that a first nucleotide sequence, such as a regulatory element, is fused in frame with a second nucleotide sequence so as to afford a faithful transcription of the entire nucleotide sequence, which upon translation yields the desired protein.

The term "immunoreactivity" and related terms refers to the ability of antibodies and fragments thereof to bind to particular regions (antigens) presented by polypeptides and proteins, presented to the antibodies either as immunogens or targets. Typically, the binding affinity of the antibodies for their antigen is in the range 10^5 to 10^{11}, with higher affinities being preferred.

The term "specific immunoreactivity" refers to the ability of antibodies and fragments thereof to bind to particular regions (antigens) presented by polypeptides and proteins, presented to the antibodies either as immunogens or targets and not to unrelated antigens. For example, an antibody with specific immunoreactivity to actin will bind actin but would not bind another protein, such as a polymerase, which do not share epitopes with actin.

The term "nucleic acid molecule" refers to DNA, RNA and nucleic acid molecule analogs such as PNA and the like. PNA or "Peptide Nucleic Acid" is a nucleic acid molecule analog that has a neutral "peptide-like" backbone with nucleobases that allow the molecule to hybridize to complementary RNA or DNA with higher affinity and specificity than corresponding

oligonucleotides. PNA can be made to be more resistant to normal nucleases and are especially desirable, for example, in gene therapy. PNA is known to one of skill in the art and can be purchased or custom synthesized in numerous commercial laboratories including PerSeptive Biosystems, Inc. (Framingham, MA).

The term "modulate" means to activate or inhibit. For example, a PARG modulator may activate or inhibit PARG activity. "Modulation activity" means the amount of activation or inhibition. For example, a compound that increase PARG (or any other enzyme) activity by 10% will have a modulation activity of 10%. Conversely, a compound that decreases PARG activity by 10% will have a modulation activity of -10%.

As used herein, a given nucleotide or amino acid sequence is said to have a defined *percentage of sequence similarity* with another sequence when the two sequences differ by no more than the specified sequence similarity, including conservative substitutions, insertions, and deletions. Degenerate codons do not result in a change in amino acid upon translation, therefore, it is appreciated that identical amino acids can be encoded by several equivalent codons. The term "homology" and "sequence similarity" should have the same meaning for the purpose of this patent. Similarity parameters may be any generally acceptable parameter. For the purposes of this patent, percent similarity between two polymers such as nucleic acid molecules and polypeptides is preferably defined by Karlin and Altschul (*11*). The similarity algorithms of Karlin and Altschul are well known to those of skill in the art as exemplified by their adoption by the National Center for Biological Information. For nucleic acid molecule sequence searching, one desirable set of parameters would M (score for a pair of matching residues) at 5; N (score for mismatching residues) at -4; W (word length) at 11. For proteins, it is well known that some amino acids are similar and that substitution would be conservative. That is, for example, the replacement of an acidic amino acid with another acidic acid would be consider a conservative mismatch while the replacement of an acidic amino acid with a basic amino acid would be consider a more divergent mismatch. Preferably, the parameters for a desirable protein similarity determination are expressed in the sequence similarity matrix BLOSUM62 as described in Henikoff & Henikoff (*12*). Other similarity matrixes that are also preferred in the invention are PAM40, PAM120 and PAM250 as described in Altschul (*13*).

The rapid synthesis of ADP-ribose polymers that occurs in response to DNA strand breaks is accompanied by very rapid polymer turnover, indicating that PARP and PARG

activities are closely coordinated as cells respond to DNA damage. While PARP has been widely studied, information concerning structure and function relationships of PARG is much more limited. The present invention discloses the isolation of a cDNA encoding the bovine, human, murine and drosophila PARG and their deduced amino acid sequences.

The availability of PARP cDNA has allowed a number of molecular genetic approaches to study the function(s) of ADP-ribose polymer metabolism and the availability of PARG cDNA should allow the design of additional molecular genetic approaches for studying this metabolism. For example, disruption of the gene encoding PARG in mice containing a normal PARP gene will allow the determination of whether other cellular enzymes can replace PARG in the turnover of ADP-ribose polymers and/or whether development of animals will occur in the absence of PARG. Alternatively, disruption of the PARG gene in mice containing a disrupted PARP gene may provide insights for the coordinated function of PARP and PARG.

One embodiment of the invention is directed to a deoxyribonucleic acid (DNA) molecule that encodes a polypeptide having poly(ADP-ribose) glycohydrolase (PARG) activity. Preferably, the molecule is of mammalian origin, such as, for example, of human origin.

In a preferred embodiment, a DNA molecule of the invention comprises a nucleotide sequence with at least about 70% sequence similarity with a sequence shown in a sequence shown in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, and SEQ ID NO: 9. Higher degrees of sequence similarity, such as about 80%, about 90%, and about 100% are preferred. Most preferred is a DNA molecule comprising a nucleotide sequence substantially identical with any one of sequence shown in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, and SEQ ID NO: 9. It is preferred that a DNA molecule of the present invention comprises at least about 1000 nucleotides and has a nucleotide sequence with at least 80% sequence similarity with a sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, and SEQ ID NO: 9. Most preferably, the DNA molecule consist of a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, and SEQ ID NO: 9.

For a DNA molecule of the present invention based on a human PARG gene it is preferred that the molecule comprises a nucleotide sequence that shows similarity to the sequence shown in SEQ ID NO: 3 from about residue 2113 to about residue 3105. More preferably, the sequence similarity is from about residue 1240 to about residue 3105. Still more preferably, the

DNA molecule comprises a nucleotide sequence similarity to the coding sequence for the full-length hPARG as shown in SEQ ID NO: 3 from about residue 175 to about residue 3105.

A DNA molecule of the present invention affords probes and primer molecules that can be used in hybridization assays and PCR amplification. An exemplary oligonucleotide is less than about 1000 residues in length and comprises a nucleotide sequence at least about 10 residues long to ensure hybridization. Preferably, the at least about 10 residue region of the oligonucleotide is complementary to a sequence shown in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, and SEQ ID NO: 9. Typically, the oligonucleotide will be a DNA molecule, which can be labeled by any method as desired, for example, with a radiolabel, a fluorescence label, or chemi-luminescent label.

Another embodiment of the invention is directed to a nucleic acid molecule that hybridizes to in a nucleic acid blot (Southern blot, Northern blot) to a sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, or SEQ ID NO: 9 under stringent hybridization conditions. A nucleic acid blot may be made using techniques defined in Molecular Cloning, Second Edition, Sambrook et al., Cold Spring Harbor Press, Cold Spring Harbor, New York. DNA to be analyzed may be separated in agarose or acrylamide gels. The DNA may be transferred to nylon or nitrocellulose membrane using techniques known to those in the art. Stringent hybridization condition may be for example, prehybridizations 42 °C in 50% formamide, 0.25 M sodium phosphate buffer, pH 7.2, 0.25 M NaCl, 7% SDS, 1 mM EDTA for 10 hours, 100 ug denatured salmon sperm DNA, hybridization at 42 °C in 50% formamide, 0.25 M sodium phosphate buffer, 100 ug denatured salmon sperm DNA, pH 7.2, 0.25 M NaCl, 7% SDS, 1 mM EDTA, 1 ng/ml probe with a specific activity of 10^9 cpm/ug DNA, for 16 hours. The probe may comprise any contiguous sequence from SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, or SEQ ID NO: 9. Preferably, said contiguous sequence is at least about 50 bases long, more preferably, the contiguous sequence is at least about 75 bases long, such as at least about 100 bases, at least about 200 bases long or at least about 300 bases long. The complete sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, or SEQ ID NO: 9. Methods of labeling probes to with radioactive labels are known to those of skill in the art.

Method of washing after stringent hybridization are known. A stringent washing may comprise, for example, two washes at in 2X SSC, 0.1% SDS for 15 minutes each at room

temperature; two washes in 0.2 X SSC, 0.1% SDS for 15 minutes each at room temperature; and a final three washes in 0.2 X SSC, 0.1% SDS for 15 minutes each at 60°C. The final wash may be increased in temperature for reduced background. For example, the final wash may be a final three washes in 0.2 X SSC, 0.1% SDS for 15 minutes each at 65°C or a final three washes in 0.2 X SSC, 0.1% SDS for 15 minutes each at 68°C.

If a radioactive probe is used, hybridization may be monitored using known techniques such as autoradiogram or a two dimensional measurement of radioactivity.

An anti-sense oligonucleotide is also afforded by the present invention. The anti-sense molecule is typically less than about 1000 residues in length to ensure ease of synthesis, and hybridizes to an RNA molecule, e.g., messenger RNA, which has at least 70% sequence similarity with a sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9. Preferably, the anti-sense molecule is at least about 10 nucleotides in length to ensure hybridization with mRNA. Even more preferably, the anti-sense molecule may be at least about 15 nucleotides in length such as, for example, at least about 20 nucleotides in length; at least about 30 nucleotides in length; at least about 50 nucleotides in length; at least about 75 nucleotides in length; at least about 100 nucleotides in length; at least about 150 nucleotides in length; at least about 200 nucleotides in length; at least about 500 nucleotides in length; at least about 1000 nucleotides in length; or at least about 1500 nucleotides in length. It is also preferred that the molecule has a ribozyme activity so that it can degrade the mRNA that it binds to.

An antisense oligonucleotide may be used therapeutically to inhibit translation of mRNA encoding PARG. Synthetic antisense oligonucleotides may be produced, for example, in a commercially available oligonucleotide synthesizer. This invention provides a means to therapeutically alter levels of expression of a human or other mammalian PARG by the use of a synthetic antisense oligonucleotide drug that inhibits translation of mRNA encoding PARG. Synthetic antisense oligonucleotides, or other antisense chemical structures designed to recognize and selectively bind to mRNA, are constructed to be complementary to portions of the nucleotide sequence shown in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7 and SEQ ID NO: 9. An antisense oligonucleotide may be designed to be stable in the blood stream for administration to patients by injection, or in laboratory cell culture conditions, for administration to cells removed from the patient. The antisense may be designed to be capable of passing through cell membranes in order to enter the cytoplasm and nucleus of the cell by virtue of

physical and chemical properties of the antisense oligonucleotide which render it capable of passing through cell membranes (e.g., by designing small, hydrophobic antisense oligonucleotide chemical structures) or by virtue of specific transport systems in the cell which recognize and transport the antisense oligonucleotide into the cell. In addition, the antisense oligonucleotide can be designed for administration only to certain selected cell populations by targeting the antisense oligonucleotide to be recognized by specific cellular uptake mechanisms which bind and take up the antisense oligonucleotide only within certain selected cell populations. For example, the antisense oligonucleotide may be designed to bind to transporter found only in a certain cell type, as discussed above. The antisense oligonucleotide may be designed to inactivate the PARG mRNA by (1) binding to the PARG mRNA and thus inducing degradation of the mRNA by intrinsic cellular mechanisms such as RNase I digestion, (2) by inhibiting translation of the mRNA target by interfering with the binding of translation-regulating factors or of ribosomes, or (3) by inclusion of other chemical structures, such as ribozyme sequences or reactive chemical groups, which either degrade or chemically modify the target mRNA. Synthetic antisense oligonucleotide drugs have been shown to be capable of the properties described above when directed against mRNA targets (*14*). In addition, coupling of ribozymes to antisense oligonucleotides is a promising strategy for inactivating target mRNA (*15*). In this manner, an antisense oligonucleotide directed to PARG may serve as a therapy to reduce PARG expression in particular target cells of a patient and in any clinical condition that may benefit from reduced expression of PARG.

It is known by those in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding PARG, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates a nucleic acid molecule that encodes a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO:6, SEQ ID NO: 8 and SEQ ID NO: 10. The invention contemplates each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices that would encode the oligopeptides disclosed herein. These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring PARG and all such variants are to be considered as being specifically disclosed.

Although nucleic acid molecules which encode PARG and its variants preferably hybridizes under high stringency conditions to the nucleotide sequence of the naturally occurring PARG gene under appropriate conditions of stringency, it may be advantageous to produce nucleotide sequences encoding PARG or its derivatives possessing a substantially different codon usage. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding PARG and its derivatives and variants without altering the produced amino acid sequence include the production of RNA transcripts having more desirable properties, such as greater half-life, than transcripts produced from the naturally occurring sequence.

In order to express a biologically active or immunologically active PARG, the nucleic acid molecule encoding PARG or functional equivalents, may be inserted into appropriate expression vector, such as, for example a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Thus, another aspect of the present invention is an expression vector comprising a regulatory sequence operatively linked to nucleic acid molecule comprising a nucleotide sequence disclosed herein. For example, an expression vector can contain a nucleotide sequence at least about 1000 base pairs in length, which has at least about 70%, about 80%, or higher, sequence similarity with a sequence shown in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, and SEQ ID NO: 9.

Methods that are known to those skilled in the art may be used to construct expression vectors containing sequences encoding PARG and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination.

A variety of expression vector/host systems may be utilized to contain and express sequences encoding PARG. These include, for example, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus), insects infected with virus expression vectors (e.g., fall army worm infected with baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus; TMV) or with bacterial expression

vectors (e.g., Ti or bacterial plasmids); or animal cell systems. The invention is not limited by the host cell employed.

Prokaryotic expression systems are commercially available from a number of suppliers worldwide. Prokaryotic expression vectors provide a convenient system to synthesize proteins. If it is desired to express a protein with characteristics such as immunogenic properties, 3D conformation, and other features exhibited by authentic PARG, the protein may be expressed in an eukaryotic protein expression system. The eukaryotic expression systems are numerous and include mammalian, amphibian, plant, insect, and yeast expression systems.

Yeast hosts that can be used for expression include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pichia pastoris*, *Hansela polymorpha*, *Kluyveromyces lactis*, and *Yarrowia lipolytica*. Yeast hosts offer the advantages of rapid growth on inexpensive minimal media and ease in large-scale production using bioreactors. Another advantage of yeast is the ability to direct expression to cytoplasmic localization or for extracellular export.

Most yeast vectors for protein expression are derivatives of the S. cerevisiae 2μ (two micron) plasmid. Yeast vectors include pYES and pEST from Stratagene (La Jolla, CA). Constitutive gene expression by the yeast plasmid cassette can be mediated by well known promoters such as the glyceraldehyde-3-phosphate dehydrogenase promoter (TDH3); the triose phosphate isomerase promoter (TPI1); the phosphoglycerate isomerase promoter (PGK1); the alcohol dehydrogenase isozyme II (ADH2) gene promoter; GAL1 and GAL10 promoters; the metallothionein promoter from the CUP1 gene (induced by copper sulfate); and the PHO5 promoter (induced by phosphate limitation). Proper termination of yeast transcripts is known to those in the art. Termination signals may include the MF-alpha-1, TPI1, CYC1, and PGK1 genes. These termination signals may be spliced onto the 3' end of the insert to provide proper termination.

Insect expression systems include baculovirus based vectors designed to express foreign proteins in a number of insect hosts and insect cell line hosts. Insect and insect cell lines may be of Drosophila melanogaster, Aedes albopictus, Spodoptera frugiperda, and Bombyx mori origin. Numerous expression systems comprising cells, vectors, hosts and the like can be purchased from a variety of commercial sources.

The control elements or regulatory sequences necessary for the proper expression of the insert, in this case PARG, may comprises promoters, enhancers (including both proximal and

distal control elements) which interact with the host proteins to carry out transcription and translation. Such elements may vary in their strength and specificity and are known to those in the art. Depending on the vectors system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, the LacZ promoter may be used in a bacterial cell; the baculovirus polyhedrin promoter may be used in an insect cell; plant promoters such as heat shock promoters, and storage protein promoters, plant virus promoters and the like may be used in a plant cell. In a mammalian cell expression system, an SV40 promoter or EBV promoter may be used.

Methods and protocols for both prokaryotic and eukaryotic expression systems are generally known to those in the art. Further, the cells, vectors, growth medium may be purchased from commercial suppliers. The catalogs and product literature of commercial suppliers provide detailed protocols to enable the expression of proteins in prokaryotic and eukaryotic systems including bacterial, yeast, insect, insect cell, and mammalian cell systems. The product literature and catalogs of Clontech (Palo Alto, CA), Invitrogen (Carlsbad, CA), Life Technologies (Rockville, MD), Novagen (Madison, WI), Pharmigen (San Diego, CA), Quantum Biotechnologies (Montreal, Quebec, Canada), and Stratagene (La Jolla, CA) are incorporated herein by reference.

A further aspect of the invention is isolated proteins and protein fragments having poly(ADP-ribose) glycohydrolase (PARG) activity. Such a protein can comprise an amino acid sequence with sequence similarity of at least about 70%, about 80% or higher to a sequence shown SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, and SEQ ID NO: 10. For example, the full-length bovine PARG has a molecular weight greater than about 100 kDa, thereby distinguishing it from previously known PARGs. The protein may be purified, for example, from cell lysates using the antibodies of the invention. The purification may be through an antibody column.

PARG polypeptides are another aspect of the invention. Polypeptides of PARG may be used, for example, to generate antibodies in an immunogenic procedure. To be effective it is preferred that the polypeptides are at least about 6 amino acid residues in length, such as for example, at least about 10 amino acids in length, at least about 20 amino acids in length, at least about 30 amino acids in length, at least about 50 amino acids in length, at least about 75 amino acids in length, at least about 100 amino acids in length, at least about 150 amino acids in length,

at least about 200 amino acids in length, or at least about 400 amino acids in length. In one embodiment, the polypeptide has a molecular weight less than about 65 kDa and with at least about 80% sequence similarity with a sequence shown in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, and SEQ ID NO: 10. The polypeptide may consist of the sequence set forth in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, or SEQ ID NO: 10.

The polypeptide of the invention may be conjugated to a larger molecule, such as, for example, keyhole lymphet hemocyanin (KLH), to increase the immunogenicity of the polypeptide. The increased immunogenicity of the polypeptide will, in turn, increase the yield of antibody. Preferably, the polypeptide has a molecular weight less than about 40 kDa and with at least about 90% sequence similarity with a sequence shown in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, and SEQ ID NO: 10. The polypeptide can also be used in a wide variety of assays, e.g., as a competitor of antigen in a liquid sample in an antibody-based assay. Therefore, it is preferred that the polypeptide has poly(ADP-ribose) glycohydrolase (PARG) activity. A particularly preferred polypeptide is of human origin and comprises an amino acid sequence substantially identical with SEQ ID NO: 4 from about residue 647 to about residue 977 - the C terminus catalytic region of the enzyme. Longer sequences more inclusive of the natural molecule are of course also contemplated.

The invention also encompasses PARG variants and alleles. A preferred PARG variant is one having at least 80% and more preferably at least 90% amino acid similarity to the amino acid of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 and SEQ ID NO: 10 and which retains at least one biological, immunological or other functional characteristic or activity of PARG. A most preferred PARG variant is one having at least 95% amino sequence similarity or identity to human PARG (SEQ ID NO: 3).

Antibodies to PARG may be generated using numerous established methods that are well known in the art. One example of such a method is described in the Examples. Generated antibodies may include, for example, polyclonal, monoclonal, chimeric, single chain, Fab fragments, Fab' fragments, Fab'(2) fragments, and fragments produced by a FAB expression library. Humanized antibodies and single chain antibodies may also be produced after the amino acid sequence of effective antibodies are determined.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with PARG or any fragment or oligopeptide

thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, for example, Freund's mineral gels such as aluminum hydroxide, and surface-active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. Among adjuvant used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to PARG have an amino acid sequence consisting of at least five amino acids and more preferably at least about 10 amino acids, such as for example about 20 amino acids or about 40 amino acids. It is also preferable that they are identical to a portion of the amino acid sequence of the natural PARG. Short stretches of PARG amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and antibodies may be produced against the chimeric molecule.

Antibodies may be produced by inducing *in vivo* production in the lymphocyte population of a living animal or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in published procedures (*16*).

Antibody fragments that contain specific binding sites for PARG may be generated. For example, such fragments include the $F(ab')_2$ fragment, Fab fragment, Fab' fragment which can be produced by enzymatic digestion of the antibody molecule. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W.D. (1989) Science 254, 1275-1281).

Therapeutic Methods

A method of preventing, treating, or ameliorating a disease condition in a patient, which disease state is affected by the level of PARG expression is also contemplated. This method entails administering a therapeutically effective amount of a poly(ADP-ribose) glycohydrolase (PARG) inhibitor or activator to the individual. Particularly, implicated disease states are neoplastic disorder, myocardial infarction, vascular stroke and neurodegenerative disorders.

In one embodiment, antisense oligonucleotides for PARG may be used alone or in combination with other chemotherapeutic agents to treat neoplastic disorder. The anti-sense oligo is designed to hybridize *in vivo* to messenger RNA expressed by the organism. The use of anti-sense molecules in a therapeutic setting is described, for example, by S. Agrawal,

Antisense Therapeutics, Humana Press. Currently favored protocols call for the oligo to have ribozyme activity in an effort to degrade the mRNA. These methods are described, for example, in Therapeutic Application of Ribozymes, K. Scanlon, ed., Humana Press. Therefore, in one embodiment, an antagonist of PARG may be administered to a subject to prevent or treat neoplastic disorder.

PARG levels may be enhanced to suppress DNA repair and increase a cell's susceptibility to chemotherapy drugs. Therefore, in another embodiment, an PARG enhancer is administered to a subject along with a chemotherapeutic drug as a treatment for neoplastic disorder.

Neoplastic disorders that can be treated by PARG elevation and chemotherapy include benign and malignant neoplasm such as, for example, adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, hyperplasia and hypertrophy. Neoplastic disorders may include, in particular, neoplastic disorders of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. For the purposes of this invention, a neoplastic disorder is any new and abnormal growth; specifically a new growth of tissue in which the growth is uncontrolled and progressive. Malignant cancer is a subset of neoplastic disorders which show a greater degree of anaplasia and have the properties of invasion and metastasis.

The synthesis of effective anti-sense inhibitors is known. Numerous approaches have been previously described and generally involve altering the backbone of the polynucleotide to increase its stability in-vivo. Exemplary oligonucleotides and methods of synthesis are described in U.S. Patents 5,661,134; 5,635,488; and 5,599,797 (phosphorothioate linkages), U.S. Patents 5,587,469 and 5,459,255 (N-2 substituted purines), U.S. Patent 5,539,083 (peptide nucleic acids) and U.S. Patents 5,629,152; 5,623,070; and 5,610,289 (miscellaneous approaches). The disclosures of each of these references are incorporated herein by reference.

Significantly, the present invention discloses a method of identifying an agent that inhibits or activates poly(ADP-ribose) glycohydrolase (PARG) activity. Such method comprises (i) providing a liquid medium that contains a polypeptide of the present invention; (ii) contacting the polypeptide with a candidate agent, in the presence of a reference compound having affinity for the polypeptide, under predetermined assay conditions; and (iii) determining the affinity of

the candidate agent for the polypeptide relative to the reference compound, thereby determining the inhibition or activation activity of the candidate agent relative to the reference compound. These determinations can be facilitated by immobilizing the polypeptide on a solid support. Alternatively, the polypeptide can be generated *in vitro* by culturing a cell transformed with a PARG gene under conditions effective to express the polypeptide.

Combination therapies are also afforded by the present invention in which a PARG inhibitor or activator is administered in combination with a chemotherapeutic or a "clot-busting" drug. The clot-busting drug may be, for example, tissue plasminogen activator (t-PA) or streptokinase.

In some cases it may be desired to overexpress PARG in the cells of an organism in order to achieve the correct PARP/PARG balance. In this context of gene therapy, it is desired to stably transfect target cells with a vector, such as, for example, a viral or a DNA (nucleic acid) vector, so that the desired gene is overexpressed. Gene therapy vector systems and protocols are well known and are described, for example, in the Internet Book of Gene Therapy (*17*) Anti-sense and ribozyme approaches to cancer gene therapy are described in chapters 7-9 of the Internet Book of Gene Therapy, and are incorporated herein by reference. Another reference is Gene Therapy Protocols, P. Robbins, ed., Humana Press. Furthermore, gene therapy methods have advanced greatly and are well documented in numerous issued US patents. Gene therapy may be practiced, for example, by substituting a nucleic acid molecule of the invention with the nucleic acid molecule described in the methods referred to in any issued US patents directed to gene therapy (*18*).

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

Diagnostic Methods

Methods of genotyping an individual for a mutant PARG allele are also afforded by the present invention. A number of protocols are available for identifying a mutant allele as described herein once the nucleotide sequence encoding PARG is known. Some exemplary methods are restriction fragment length polymorphism (RFLP), allele-specific PCR (ASPCR)

and single-strand conformation polymorphism (SSCP). Armed with this information, the genetic susceptibility of an individual to an above-mentioned disease condition can be assessed.

An allele-specific method for identifying point mutations by differential PCR amplication is described by (*19*). A non-electrophoretic method of genotyping with allele-specific PCR employs a dye specific for double-stranded DNA (*20*). A method of detecting mutations referred to as single-stranded conformation polymorphism (SSCP) is presently widely employed (*21*). A hybrid of SSCP and Sanger dideoxy sequencing, called dideoxy fingerprinting (ddF) has recently been described (*22*).

Other methods of identifying allelic mutations are known to the skilled artisan. Probably the most commonly used method of genotyping is restriction fragment length polymorphism (RFLP) (*23*), which is employs one or more restriction enzymes to identify mutant alleles occurring within a restriction site. This method has been used extensively in forensic applications and is employed commercially by such companies as Helix Biotech, Inc. Reliagene Technologies, Inc. and GenTest Laboratories, Inc. Accordingly, an instant mutant PARG allele can be detected by RFLP methods, optionally by one of these commercial entities. The above methods are most effective in the detection of homozygotes for the defective allele.

An RFLP method of identifying a mutant PARG allele in an individual entails: (i) obtaining genomic material from the individual; (ii) digesting the genomic material with a restriction enzyme having a recognition site inclusive of the mutant allele; (iii) fractionating the restriction fragments obtained from the digestion, e.g., by electrophoresis; and (iv) comparing the fractionation pattern with that obtained for a normal allele, thereby determining the presence or absence of the mutant allele.

An ASPCR method of identifying a mutant PARG allele in an individual entails: (i) hybridizing an oligonucleotide with genomic material from the individual; (ii) attempting to extend the oligonucleotide using PCR amplification; and (iii) determining the degree to which extension occurs, thereby determining the presence or absence of the mutant allele. In this method, it is preferred that the oligonucleotide hybridizes under predetermined hybridization conditions to a region immediately 5' of a predetermined mutation site in the PARG allele with the 3' terminus of the oligonucleotide complementary to an unmutated PARG allele. In these protocols, the PCR extension reaction is generally attempted at a temperature above about 50°C, more preferably above about 60°C.

A variety of protocols including ELISA, RIA and FACS for measuring PARG levels are known in the art and provide a basis for diagnosing altered or abnormal levels of PARG expression. Normal or standard values for PARG expression may be established by combining body fluids and tissue biopsies from normal mammalian subjects, rupturing the cells or permeating the cells, combining the cells with antibody under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods but preferably by photometric means. Quantities of PARG expressed in subject, control, and disease sample are compared to standard values to determine between normal, reduced or enhanced levels of PARG.

A still further aspect of the invention pertains to an antibody immunoreactive with a polypeptide of the present invention. Preferably the antibodies are specifically immunoreactive with the polypeptides of this invention such as, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, and SEQ ID NO: 10. Frequently it is desired to label the antibody, e.g., with a radiolabel, fluorescent or epitope label, to permit visualizing the antibody. Thus, antibodies immunoreactive with the PARG of this invention are afforded, which can be used to study features of PARG heterogeneity and possible modes of regulation. The high degree of sequence similarity between bovine PARG, human PARG and murine PARG permits eliciting antibodies to PARG of one species, which are found to be cross-reactive with PARGs from other organisms. These antibodies are valuable in characterizing PARG in-vivo under defined physiological conditions in many different organisms.

Accordingly, a method of detecting a polypeptide having PARG activity, for example, a diagnostic assay, entails: (i) contacting the polypeptide with an aforementioned antibody of the invention; and (ii) determining whether the antibody immunoreacts with the polypeptide. Binding can be ascertained in an sandwich assay, as is well known, due to the ability of the antibodies to immunoreact with an epitope of PARG. Preferably, monoclonal antibodies, such as those prepared by the method of Kohler and Milstein (*24*) and labeled antigens effective in competing with the polypeptide, are employed. Exemplary assays are disclosed in U.S. Patent 4,375,110, the disclosure of which is incorporated herein by reference.

The present invention includes immunoreactive fragments of a PARG enzyme. Immunoreactive fragments can be fragments that can elicit an immune response that recognizes a PARG enzyme. Alternatively, immunoreactive fragments can be fragments that are

specifically bound by an antibody that specifically binds a PARG enzyme. Any of variety of methods may be employed in order to identify contiguous peptide fragments of a PARG enzyme that comprise immunoreactive sequences. PARG enzymes may be fractionated by proteases, cyanogen bromide, etc. and the resultant fragments assessed for their capacity to specifically bind anti-PARG antibodies.

In an alternative embodiment, one or more synthetic peptides may be prepared in order to locate contiguous amino acid sequences that are immunoreactive. The peptides may have a sequence that includes a series of contiguous amino acids that are identical to a series of contiguous amino acids of a PARG enzyme. The peptides may be of about six amino acids to about 500 amino acids in length. The peptide may also include sequences that are not identical to sequences of a PARG so long as it includes at least about six contiguous amino acids that are identical to about six contiguous amino acids of a PARG enzyme. In a preferred embodiment, the peptide will be about 50 amino acids in length. In other preferred embodiments the length of the peptide may be from about six amino acids to about 30 amino acids.

The peptides of the present invention may comprise amino acid sequences that elicit antibodies that specifically bind to the peptide or to a PARG enzyme. Alternatively, the peptides may contain sequences that are specifically bound by anti-PARG antibodies. Peptides that are bound by anti-PARG antibodies may identified through the use of Epitope Scanning™ strategy (Cambridge Research Biochemicals, Inc.). Thus, the linear sequence of amino acids of a particular PARG enzyme is used to construct a set of peptides of defined length which overlap other members of the set by one or more residues. The peptides may be any length; however, lengths of from about 6 to 25 amino acids are preferred. In selecting the length, a general consideration is that antibodies that recognize linear native epitopes constitute approximately 60-70% of the anti-protein antibody population (*25*).

The number of overlapping amino acids will generally be more than half of the length of the peptides. That is, if the peptides are about 20 amino acids long, the overlap may be 11 or more amino acids long. In preferred embodiments, each peptide will be selected such that the number of overlapping amino acid residues in adjacent peptides is from about (n-1) to (n-3), where "n" is the number of amino acids in the peptide. An overlap of (n-1) is particularly preferred. Thus, in a particularly preferred embodiment, a first peptide may have the amino acid sequence of residues 1-10 of a PARG enzyme, a second peptide may have the amino acid

sequence of residues 2-11 of the same PARG enzyme, a third peptide may have the sequence of residues 3-12 of the same PARG enzyme and so on until the entire sequence of the PARG enzyme has been synthesized in fragments.

The peptides may be synthesized using any means known to those of skill in the art. In a preferred embodiment, the peptide will be synthesized using an automated synthesizer such as a multipin peptide synthesis system. Such systems or peptides synthesis services are commercially available from suitable providers known to those skilled in the art.

To identify suitable peptides, each peptide is introduced into a well of a microtiter plate, and assayed for its ability to bind to antibodies elicited by a PARG enzyme. Such assays may be conducted in various ways known to those skilled in the art. One suitable assay is conducted by immobilizing a peptide on the surface of a well and then contacting the peptide with a solution containing an anti-PARG antibody. After washing, the well is contacted with a labeled antibody that specifically binds to the anti-PARG antibody. Thus, the presence of label in the well indicates that the anti-PARG antibody bound to the immobilized peptide. Another preferred method of determining the ability of the peptide to be specifically recognized by anti-PARG antibodies is a competitive ELISA.

Once a particular peptide has been found to bind to anti-PARG antibodies, the peptide can be used to elicit monospecific antibodies. By immunizing an experimental animal with a single peptide containing a single antigenic determinant, the antibodies elicited will all specifically bind to the same antigenic determinant even though the antibodies are not monoclonal.

Where desired, the peptides can be modified to increase their immunogenicity. Thus, they may be modified to contain an amino-terminal and/or a carboxyl-terminal cysteine or lysine residue with or without spacer arms. The peptides may be conjugated to carriers such as bovine serum albumin, ovalbumin, human serum albumin, KLH (keyhole limpet hemocyanin) or tetanus toxoid. The use of human serum albumin is preferred over ovalbumin or bovine serum.

The peptides, alone or conjugated to a carrier, may be themselves capable of eliciting an antibody response when administered to an experimental animal. Alternatively, the peptides, alone or conjugated to a carrier, may be administered in conjunction with an adjuvant. Those skilled in the art will understand that a variety of materials may function as adjuvants. Examples of possible adjuvants include, but are not limited to, Freund's complete adjuvant, Freund's

incomplete adjuvant, lipopolysaccharide (LPS) and the like. Any material that increases the immune response to a fragment of a PARG enzyme may be used as an adjuvant.

The ability to produce large amounts of active PARG enzyme permits, for the first time, the large scale screening of chemical libraries for molecules capable of inhibiting or activating PARG enzymatic activity. The screening may be conducted using any assay for PARG known to those skilled in the art. In a preferred embodiment, the screen may be conducted using the TLC based assay described by Ménard, *et al.* (*26*). A known amount of PARG will be incubated under standardized conditions with [^{32}P]-poly(ADPR) in the presence of inhibitor or activator. After an appropriate period of time, the reaction will be stopped and the reaction mixture separated on PEI-F cellulose TLC plates. The TLC plates may be developed in an appropriate solvent system such as methanol followed by 0.3N LiCl. The amount of ADPR released in the reaction will be quantified and the effect of the inhibitor or activator on enzymatic activity will be determined. Typical reaction conditions are 50 mM potassium phosphate (pH 7.5) at 37 °C in the presence of 25 μM [^{32}P]-poly(ADPR). The concentration of the inhibitor or activator can be varied as necessary to determine the K_i value of the inhibitor or activator according to standard procedures.

Another embodiment of the invention is directed to a method of altering the response of the cell to a genotoxic stress by modulating the concentration of ADPR polymers. As discussed above, the metabolism of ADPR polymers is critical in determining the fate of cells subjected to genotoxic stress. The modulation can be either an increase or a decrease in the concentration of the polymers. In one embodiment of the present invention, the concentration of ADPR polymers can be decreased by the use of a gene therapy vector expressing a high level of PARG. In another embodiment of the present invention, the concentration of polymers can be increased by inhibiting the enzymatic activity of the PARG enzyme by the addition of inhibitors or activators identified as described above. Alternatively, the concentration of ADPR polymers can be increased by interfering with the endogenous expression of PARG enzymes using antisense oligonucleotide technology.

Knowledge of the nucleotide sequence of the PARG gene permits the preparation of antisense therapeutics containing sequences complimentary to the mRNA of PARG gene. The preparation and delivery of antisense therapeutics is well known to those skilled in the art. For example, antisense therapeutics have been used to treat neoplastic disorder as exemplified by

Smith, U.S. Patent No. 5,248,671, specifically incorporated herein by reference. Additional examples of antisense therapeutics are provided by Miller, U.S. Patents No. 4,511,713 and 4,757,055, specifically incorporated herein by reference.

In the present invention, an oligonucleotide having a sequence complimentary to the mRNA of the PARG gene will be prepared. Such an oligonucleotide is said to be an antisense oligonucleotide with respect to the PARG gene. The oligonucleotide may be RNA or DNA or a may contain both RNA and DNA portions. The oligonucleotide may contain modified bonds so as to enhance the stability of the oligonucleotide and render it more resistant to the action of cellular nucleases. For example, the oligonucleotide may be constructed with phosphorothioate nucleotides, phosphonate nucleotides and other types of modified nucleotides known to those skilled in the art. The structure of the oligonucleotide may be altered so as to include other types of bonds that do not naturally occur in oligonucleotides. For example, adjacent nucleosides might be joined using linear alkyl chains, peptide bonds or other types of structures. The only limitation is that the resulting oligonucleotide remains capable of hybridizing to the target PARG mRNA.

The antisense oligonucleotides may be delivered by any means customarily used in the art. For example, the oligonucleotide may be delivered in neutral liposomes, cationic liposomes or by ballistic high speed injection. Alternatively the DNA sequence encoding the antisense oligonucleotide may be inserted into a gene vector and the vector may be introduced into target cells. The vector may be any type of gene therapy vector known to those skilled in the art. Preferred embodiments include, plasmid vectors and viral vectors. Viral vectors are seen to include those vectors customarily used for gene therapy applications including, but not limited to, retroviral vectors, vaccinia virus vectors, herpes virus vectors, adenovirus vectors and adeno-associated virus vectors. Upon introduction of the vector into target cells, the vector will direct expression of a nucleic acid molecule comprising the appropriate sequence to hybridize with the mRNA encoding a PARG enzyme. In a preferred embodiment, introduction of the vector into the target cell will result in the production of an RNA molecule that hybridizes with the mRNA of a PARG enzyme and also includes one or more additional RNA sequences capable of functioning as a ribozyme. The ribozyme portion of the molecule will cause the cleavage of the mRNA encoding the PARG enzyme thereby preventing the production of PARG.

Therapeutics of this type may be used to treat a wide variety of conditions. In one embodiment, an antisense therapeutic will be used to treat neoplastic disorder. In a preferred embodiment, an antisense therapeutic of the present invention will be delivered in combination with a currently known chemotherapeutic agent. In general, chemotherapeutic agents function by disrupting the integrity of DNA in target cells. Since the recovery of a cell from such DNA disruption is highly dependent upon the normal ADPR polymer metabolism, the presence of the antisense therapeutic will have the effect of chemosensitizing the neoplastic cells by disturbing the ratio PARG and PARP.

In another preferred embodiment, the antisense oligonucleotides of the present invention may be used to treat a variety of conditions caused by genotoxic oxidative stress. Examples include cardiac disorders, neuronal disorders, reperfusion injury, neurotoxicity, Alzheimer's disease, Huntington's disease and Parkinson's disease. It has been shown that inhibition of ADPR polymer synthesis provides protection against cellular damage caused by nitric oxide injury. Zhang, *et al*,. U. S. Patent No. 5,587,384, specifically incorporated herein by reference, teach that decreasing the amount of ADPR polymers formed can result in protection against nitric oxide induced neurotoxicity. As discussed above, decreasing the amount of ADPR polymers in the cell can be accomplished by the introduction of gene therapy vector expressing PARG, thus, the present invention can be used to treat neurodegenerative conditions resulting from oxidative stress.

CONCLUSION

The synthesis and rapid turnover of ADP-ribose polymers is an immediate cellular response to DNA damage. Reported here is the isolation and characterization of cDNAs encoding various poly(ADP-ribose) glycohydrolase (PARG) enzymes responsible for ADP-ribose polymer turnover. PARG was isolated from bovine thymus, yielding a protein of approximately 59 kDa. Based on the sequence of oligopeptides derived from the enzyme, polymerase chain reaction products and partial cDNA clones were isolated and used to construct a putative full-length cDNA. The cDNA of approximately 4.1 kb pairs predicts expression of a protein of approximately 111 kDa, nearly twice the size of the isolated protein. A single transcript of approximately 4.3 kb pairs is detected in bovine kidney poly(A)$^+$ RNA, consistent with expression of a protein of 111 kDa. Expression of the cDNA in *Escherichia coli* results in

an enzymatically active protein of 111 kDa and an active fragment of 59 kDa. Analysis of restriction endonuclease fragments from bovine DNA by Southern hybridization indicate that PARG is encoded by a single copy gene. Taken together, the results indicate that previous reports of multiple PARGs can be explained by proteolysis of an 111-kDa enzyme. The deduced amino acid sequence of the bovine PARG shares little or no sequence similarity with differing types of known proteins; however, it contains a putative bipartite nuclear location signal as would be predicted for a nuclear protein. The availability of cDNA clones for PARG should facilitate structure-function studies of the enzyme and its involvement in cellular responses to genomic damage.

Other embodiments and advantages of the invention are set forth, in part, in the description that follows and, in part, will be obvious from this description and may be learned from practice of the invention.

EXAMPLES

Example 1 Purification of bovine PARG.

PARG was purified from bovine thymus tissue (Pel-Freez, Rogers, AK) by modifications of previously published procedures (*27*). The enzyme was isolated up to the polyethylene glycol (PEG)-6,000 fractionation step as described previously (*28*). However, DNA-agarose and heparin-Sepharose chromatographic steps used previously were omitted, and the PEG-6,000 fraction was applied directly to an affinity matrix of poly(ADP-ribose)-dihydroxyboronyl-Sepharose (PADPR DHB-Sepharose). The active fractions eluted from PADPR DHB-Sepharose (25 ml) were pooled, placed in dialysis tubing, concentrated against dry PEG-20,000 to approximately 12 ml, and dialyzed against 2 liters of 20 mM potassium phosphate buffer, pH 8.0, 0.1% Triton X-100, 5 mM β-mercaptoethanol, 0.1 mM thioglycolic acid, 0.4 M KC1 (buffer A). The sample was loaded onto a 1.0 x 11-cm Toyopearl AF-Red (Supelco) column, and PARG was eluted with an 80-ml linear gradient of 0.4-2 M KC1 in buffer A. The active fractions, eluting at approximately 1.25 M KC1, were pooled, placed in dialysis tubing, concentrated against solid sucrose to approximately 9 ml, and dialyzed against 20 mM potassium phosphate buffer, pH 7.2, 0.75 M KC1, 0.1% Triton X-100, 10% glycerol, 5 mM β-mercaptoethanol, 0.1 mM thioglycolic acid. PARG activity was determined as described by Ménard and Poirier (*29*), and protein content was determined by the method of Bradford (*30*). The final preparation was quantified

by SDS-PAGE (*31*) and Coomassie Blue staining to compare the intensity of the protein band with a known amount of bovine serum albumin (*32*).

The purification procedure for the bovine thymus PARG summarized in Table 1 is typical for results obtained from six separate preparations of the enzyme. Purification from 500 g of bovine thymus achieved approximately 50,000-fold purification and yielded approximately 20 µg of purified protein. An aliquot of the purified enzyme was precipitated with trichloroacetic acid, washed with acetone, resuspended in SDS-PAGE sample buffer, separated on a 10% SDS-PAGE gel, and stained with Coomassie Blue. Analysis of the final preparation of SDS-PAGE revealed that more than 95% of the protein migrated at an apparent molecular mass of approximately 59 kDa (Figure 2). In Figure 2, an aliquot of the purified enzyme was precipitated by TCA, washed with acetone, resuspended in SDS-PAGE sample buffer, separated on a 10% SDS-PAGE gel and stained with Coomassie blue. The positions of molecular weight marker proteins are shown.

Table 1

Purification of PARG from bovine thymus

Step	Protein	Total activity	Specific activity	Yield	Purification
	mg	*units*	*units/mg protein*	*%*	*-fold*
Crude extract	27,800	57,400	2.06	100	1.0
Protamine sulfate	12,500	58,000	4.64	101	2.3
Ammonium sulfate	4,480	30,000	6.70	52	3.3
CM-Sepharose	171	19,100	112	33	55
PEG 6000	23.0	7,530	327	13	160
PADPR-DHB-Sepharose	1.30	6,730	5,180	12	2,500
Toyopearl AF-Red	0.023	2,260	98,300	4	48,000

Example 2 Peptide Sequencing.

Prior to proteolytic fragmentation, the purified bPARG (40 μg in 100 μl of 0.4 M ammonium bicarbonate buffer, pH 8.0, 8 M urea) was incubated in a final concentration of 2.2 mM dithiothreitol at 56°C for 15 min. Iodoacetamide was added to a final concentration of 2.0 mM, and the sample was incubated at 25°C for 15 min. After dilution with an equal volume of water, 1.5 units of immobolized L-1-tosylamido-2-phenylethyl chloromethyl ketone-treated trypsin (Pierce Chemical, Rockford, IL) was added, and the sample was incubated at 37°C for 18 h with gentle rotary shaking. Finally, the mixture was subjected to centrifugation at 16,000 x *g* for 5 min to separate the tryptic fragments from the immobolized trypsin. The tryptic fragments were adjusted to 0.05% in trifluoroacetic acid and separated on a 4.6 mm x 25 cm, Microsorb MV, C_4 reversed-phase HPLC column (Rainin) eluted with an 80-min linear gradient from 4 to 44% acetonitrile in 0.05% trifluoroacetic acid. Four oligopeptide fractions, with approximate elution times of 61, 63, 68, and 75 min, were selected for peptide sequence analysis by the Edman degradation method. Amino acid sequence data of four oligopeptides, designated by their approximate HPLC elution times from the reversed-phase column, are shown in Table II.

Table II

Amino acid sequence of oligopeptides derived from bPARG

Oligopeptide	Amino Acid Sequence			SEQ ID NO:
	10	20	30	
68	LFTEVLDHNE	CLIITGTEQY	SEYTGYAETY R	SEQ ID NO: 11
63	AYCGFLRPGV	SSENLSAVAT	GNXGCGAFG	SEQ ID NO: 12
61	FLINPELIVS	R		SEQ ID NO: 13
75	IALXLPNIXT	QPIPLL		SEQ ID NO: 14

Example 3 cDNA Cloning.

To obtain cDNA clones encoding bovine PARG, PCR amplification experiments were followed by the screening of two different bovine cDNA libraries. Figure 3 depicts the alignment of the DNA sequences of two PCR products and eight λ gt11 cDNA clones used to identify the cDNA coding for bovine PARG. The two PCR products and clones 1 and 2 were

obtained from the bovine thymus cDNA library. Clones 3-8 were obtained from the bovine kidney cDNA library. The positions of restriction sites used in this study are shown, and the top diagram shows the consensus clone, denoting the relative location of the coding regions for oligopeptides, 75, 61, 68, and 63 as well as the open reading frame and noncoding regions. For each of the cDNA inserts characterized, the sequence of both strands was determined by the dideoxynucleotide chain termination method using Sequenase™ (U.S. Biochemical Corp., Cleveland, OH).

The first step leading to the isolation of cDNA clones was to synthesize two multi-degenerate 17-mer primers, GAYCAYAAYGARTGYYT (SEQ ID NO: 15) and CKRTANGTYTCNGCRTA (SEQ ID NO: 16) (where Y represents T/C, R is A/G, K is T/G, and N is A/T/C/G), based on two regions of the SEQ ID NO: 11; "DHNECL" (amino acids 7 to 12 of SEQ ID NO: 11) and "YAETYR" (amino acid 26 to amino acid 31 of SEQ ID NO: 11) (Table II). Using the multidegenerate primers and an oligo(dT)-primed bovine thymus cDNA λgt11 library BL1019b from Clontech (Palo Alto, CA), PCR amplification generated a 74-bp DNA fragment with a deduced amino acid sequence identical to the corresponding region of oligopeptide 68. Next, two specific 24-mer oligonucleotide primers, ATCATCACAGGTACTGAGCAGTAC (SEQ ID NO: 17) and GCCTGTGTATTCACTGTACTGCTC (SEQ ID NO: 18), based on the sequence of this 74-bp DNA were used in combination with λgt11 forward and reverse primers to amplify PCR products 1 and 2 from the bovine thymus library. PCR product 1 contained 231 bp of sequence including the region encoding the N-terminal region of oligopeptide 68 (SEQ ID NO: 11) and the entire sequence of oligopeptide 61 (SEQ ID NO: 13). PCR product 2 contained 757 bp, which included a sequence encoding the C-terminal region of oligopeptide 68 (SEQ ID NO: 11) and the entire sequence of oligopeptide 63 (SEQ ID NO: 12).

The sequence information obtained from PCR products 1 and 2 was used to isolate cDNA clones obtained by the screening of bovine thymus and bovine kidney cDNA libraries. A 518-bp *Eco*RI-*Hind*III fragment from PCR product 2 was used as a probe to screen approximately 1 x 10^6 independent clones from the bovine thymus library. Two positive cDNA clones (clones 1 and 2) were isolated, which overlapped PCR products 1 and 2. However, attempts to obtain clones from the bovine thymus library that contained sequence 5' to clone 2 were unsuccessful. Thus, a 231-bp *Eco*RI-*Kpn*I fragment from clone 2 was used as a probe to screen approximately

5 x 10^5 independent clones of the bovine kidney 5' stretch plus cDNA λgt11 library BL300lb (Clontech, Palo Alto, CA). Three positive cDNA clones (clones 3-5) were obtained, all of which contained sequence 5′ to clone 2. Each of these clones also contained a sequence encoding oligopeptide 75. Clones 1-5 provided multiple overlapping sequences in the 3′-terminal portion of a consensus cDNA, but additional clones were sought to obtain overlapping sequences for the 5′-terminal region. Thus, a 436-bp *Eco*RI-*Kpn*I fragment located at the 5' end of clone 3 was used as a probe to screen approximately 6 x 10^5 independent clones of the bovine kidney library. Clones 6-8 provided overlapping sequences for the 5′-terminal region. The full-length cDNA was constructed by ligating a 3.9-kb *Xba*I-*Nsi*I fragment from pWL11 (clone 1 cDNA insert in pTZ18R (*33*)) and a 3.0-kb *Nsi*I-*Xba*I fragment from pWL13 (clone 4 cDNA insert in pTZ18R). The resulting plasmid, termed pWL30, contained the 4,070-bp full-length cDNA.

Figure 3 shows an alignment of the DNA sequences of two PCR products and eight λgt11 cDNA clones used to identify the cDNA coding for bovine PARG. The two PCR products and Clones 1 and 2 were obtained from the bovine thymus cDNA library. Clones 3 through 8 were obtained from the bovine kidney cDNA library. The position of restriction sites used in this study is shown and the top diagram shows the consensus clone, denoting the relative location of the coding regions for oligopeptides 75, 61, 68, and 63 as well as the open reading frame and non coding regions.

The nucleotide sequence of cDNA coding for bovine PARG is shown in the sequence listing as SEQ ID NO: 1. The deduced amino acid sequence of the enzyme is shown in the sequence listing as SEQ ID NO: 2. The four oligopeptides sequenced from purified enzyme is within SEQ ID NO: 2. They are IALCLPNICTQPIPLLK (amino acid 601 to 617, SEQ ID NO: 2); LINPELIVSR (amino acid 761 to 770, SEQ ID NO: 2); LFTEVLDHNECLIITGTEQYSEYTGYAETYR (amino acid 771 to 801, SEQ ID NO: 2) and AYCGFLRPGV PSENLSAVAT GNWGCGAFGGDAR (amino acid 849 to 880, SEQ ID NO: 2). The combined nucleotide sequence of Clones 1 through 8 predicted a full-length cDNA clone of 4,070 bp containing 257 bp of 5′-non-coding sequence, a single open reading frame of 2,931 bp (beginning at the ATG at position 258 of SEQ ID NO: 1) and a 3′-non-coding region of 882 bp. and the deduced amino acid sequence which predicts a protein of 977 amino acids and a molecular weight of 110.8 kDa.

Example 4 Analysis of the sequence of bovine PARG.

The cDNA clone (SEQ ID NO: 1) has features typical of cDNAs that code for mammalian proteins. These include (i) an oligo A (putative poly(A)+) sequence at the 3′-end, (ii) a polyadenylation signal (AATAAA) 12 bp upstream from the oligo A sequence, (iii) a sequence of ATTTA in the 3′-untranslated region thought to play a role in selective mRNA degradation in mammalian cells (*34*), (iv) a single open reading frame, and (v) a nucleotide sequence around the first start codon commonly found at known sites of initiation of translation (*35*). The evidence that the cDNA clone constructed represents a full-length or nearly full-length clone for PARG is shown by the observation that hybridization of poly(A)+ RNA from bovine kidney cells with the cDNA showed a single band of hybridization of approximately the same size as the cDNA under stringent hybridization conditions (set forth above) (Figure 4).

The nucleotide sequence encoding bovine PARG indicates that PARG shares little or no sequence similarity with other known sequences. A search of sequence data banks has failed to reveal significant sequence similarity with any sequences coding for known proteins. A strong sequence similarity has been observed with human and rat cDNA clones that likely represent partial clones for PARG from these species. Examination of protein sequence databases such as Genbank and SwissPro also has shown that the deduced amino acid sequence of PARG lacks any sequence similarity with known proteins. However, the amino acid sequence shares a significant similarity with a protein sequence from *Caenorhabditis elegans* that may represent the PARG protein from this organism (*36*).

The deduced amino acid sequence of PARG has been examined for a number of structural motifs that can be predicted from the primary amino acid sequence. The expressed PARG protein was observed to be able to form dimers stable to SDS-PAGE conditions. In that regard, residues 871-907 show significant homologies to known leucine zipper dimerization sequences (*37*).

Another motif identified is a putative bipartite nuclear location signal (NLS) (*38*). It is interesting that PARP also contains a bipartite NLS (*39*). Figure 5 compares deduced amino acid sequences in the NLS region of the bovine PARG, and regions of putative PARG sequences from human, mouse and *C. elegans*, with the NLS region of PARP from seven different organisms. Conserved residues are noted in bold and the amino acid distances are from the amino terminal methionine residue. Abbreviations and references for the sequences shown are as follows:

bPARG, bovine PARG (SEQ ID NO: 19); *hPARG*, human PARG (SEQ ID NO: 20); *mPARG*, murine PARG (SEQ ID NO: 21); *CePARG*, *Caenorhabditis elegans* PARG (SEQ ID NO: 22); *hPARP*, human PARP (SEQ ID NO: 23; *40*); *mPARP*, murine PARP (SEQ ID NO: 24; *41*); *bPARP*, bovine PARP (SEQ ID NO: 25, *42*); *aPARP*, chicken PARP (SEQ ID NO: 26; *43*); *X1PARP*, Xenopus laevis PARP (SEQ ID NO: 27; *44*); *DmPARP*, *Drosophila melanogaster* PARP (SEQ ID NO: 28; *45*); *SpPARP*, Sarcophaga peregrina PARP (SEQ ID NO: 29; *46*). In Figure 5, conserved residues are noted in boldface type, and the amino acid distances are from the amino-terminal methionine residue. Sequence alignment of putative bipartite nuclear localization signal of bovine, human and murine PARG compared to the nuclear localization signal of PARP from different organisms. The putative NLS of PARG fulfills the criteria for bipartite NLS in that it contains conserved acidic and basic amino acid residues at two different locations each within the region of sequence similarity to the NLS of PARP (*47*).

A surprising finding was that the bovine PARG cDNA clone codes for a protein of approximately 111 kDa, which is nearly twice the size of the PARG protein isolated from bovine thymus (Figure 2). It indicates that PARG contains a protease sensitive site that, following proteolysis, yields a protein fragment of approximately 59 kDa that still retains enzymatic activity. Several pieces of evidence favor this possibility. (i) Expression of the carboxyl terminal portion of the cDNA resulted in enzymatic activity (Figure 6, bar 5). (ii) All of the oligopeptides sequenced were located in the carboxyl terminal half of the protein (Figure 3, Figure 6 and Table 2). (iii) The only protein, other than 59 kDa protein detected in the thymus preparation was approximately 111 kD (Figure 2). (iv) The PARG activity expressed in bacteria was sensitive to proteolysis, yielding a protein of approximately 56 kD (Figure 6). (v) The cleavage site in PARG is in the region of the putative NLS and the PARP NLS is located in a protease sensitive site (*48*). Taken together with the data suggesting that bovine PARG appears to be coded for by a single copy gene (Figure 7), proteolysis seems likely to explain the presence of PARG activity of molecular weight of approximately 74 kDa and 59 kDa in bovine thymus preparations (*49*). Likewise, a similar mechanism could explain previous reports of a PARG of 74 kDa isolated from nuclear fractions of guinea pig liver and human placenta (*50*) and a PARG of 59 kDa isolated from postnuclear fractions of guinea pig liver (*51*).

While proteolysis of a larger protein to yield smaller proteins retaining PARG activity seems likely to explain the size heterogeneity of PARG previously reported, it remains to be

determined if proteolysis normally occurs *in vivo* or whether it occurs during purification of the enzyme. While the results presented here show that a full-length protein can be expressed containing PARG activity (Figure 8), the molecular size of PARG *in vivo* also remains to be determined. If PARG occurs as a larger protein, an interesting possibility is that the amino terminal region may be involved in the regulation of enzymatic activity.

Example 5 Expression of bPARG in Escherichia coli.

To determine whether the isolated cDNA encoded PARG, bPARG was express using two different bacterial expression systems, the pTrcHis Xpress System™ (Invitrogen, Carlsbad, CA), in which the expressed protein contains a leader polyhistidine sequence, and the glutathione S-transferasae (GST) gene fusion system (Pharmacia Biotech Inc., Piscataway, NJ). For expression in the pTrcHis Xpress system, three different DNA fragments were amplified and inserted into the pTrcHis expression plasmid. Constructs A and B contained the entire opening reading frame of 110.8 kDa, which together with the fusion partner predicted a protein of about 115 kDa. Construct B also contained the 3'-untranslated region of the clone. Construct A, containing the cDNA sequence-3 to 2,946, was prepared by subcloning a 2.9kb *Xho*I-*Eco*RI DNA fragment amplified from pWL30 with primers WIN34 (GCTGCGGGTCTCGACGATGAGTGCGGGC) (SEQ ID NO: 30) and WIN15 (GCGTCTAGAATTCACTTGGCTCCTCAGGC) (SEQ ID NO: 31). Construct B, containing the cDNA sequence-3 to 3,813, was prepared by subcloning a 3.8-kb *Xho*I-*Eco*RI DNA fragment amplified from pWL30 with primers WIN34 (SEQ ID NO: 30) and WIN33 (CCGGAATTCGGGTTTTTTGTTAATGAAAATTTATTAAC) (SEQ ID NO: 32). Construct C, containing cDNA sequence 964-2,946, was prepared by subcloning a 2.0-kb DNA fragment amplified from pWL13 with primers WIN14 (TCAGAGCAGATGAACTCGAGCAGTCCAGG) (SEQ ID NO: 33) and WIN15 (SEQ ID NO: 31). Since the isolated PARG of approximately 59 kDa contained enzymatic activity, construct C contained only the 75-kDa carboxyl-terminal region of the PARG, which predicted a fusion protein of approximately 79 kDa.

For expression experiments of bPARG as a GST fusion protein, an insert containing the cDNA sequence from position 1138 to 2946 was prepared by subcloning a 1.8-kb *Eco*RI-*Eco*RI fragment amplified from pWL30 with the oligonucleotide CCAATTTGAAGGAGGAA-TTCCCGCCGCCACCATGAATGATGTGAATGCCAAACGACCTGGA (SEQ ID NO: 34)

and WIN15 (SEQ ID NO: 31) as primers. The resulting DNA fragment was inserted into the *Eco*RI site of the pGEX-2T expression vector, and the plasmid was used to transform *E. coli* NM522 cells.

For expression experiments, bacterial cultures were grown at 37°C in 1% Bacto-tryptone, 0.5% yeast extract, and 0.5% NaCl to a density of approximately 0.6 A_{600}/ml and were induced with 1 mM isopropyl-β-D-thiogalactoside (IPTG). Cells were collected by centrifugation, and crude extracts were prepared by sonication (10 A_{600}/ml) in 10 mM sodium phosphate buffer, pH 7.2, 150 mM NaCl, 0.5 mg/ml lysozyme, 0.1 mg/ml phenylmethylsulfonyl fluoride, 1 mM EDTA, 0.7 μg/ml pepstatin A, 0.5 μg/ml leupeptin, and 1 μg/ml aprotinin. Cell extracts were subjected to centrifugation, and the supernatant fraction was used for assay. PARG assay conditions were as described previously (*52*). Following incubations, portions of reaction mixture were analyzed by thin layer chromatography or subjected to anion exchange HPLC.

Using a thin layer chromatography assay that measures release of [^{32}P]ADP-ribose from [^{32}P]ADP-ribose polymers (*53*), PARG activity was detected in extracts from cells transformed by each of the constructs. Figure 6 shows results obtained with constructs B and C. Reaction mixtures contained approximately 15,000 cpm of [^{32}P]ADP-ribose polymers, and the cpm shown represent ADP-ribose released from the ADP-ribose polymers. Bar 1, a strain transformed by pTrcHis without an insert but induced with 1 mM IPTG for 5 h at 37°C. A strain containing construct B is shown without the addition of IPTG (bar 2) or after the addition of 1 mM IPTG for 1.5 h (bar 3) or 5 h (bar 4). A strain containing construct C 5 h after induction by IPTG is shown in the absence (bar 5) and presence (bar 6) of 167 μm ADP-hydroxymethylpyrrolidine diol (*54*). No activity was detected in cells transformed with the empty vector, but activity was detectable without induction by IPTG, indicating a leaky *lac* promoter. The addition of IPTG resulted in a time-dependent increase of up to approximately 4.5-fold in enzymatic activity. Figure 6 also shows that the enzymatic activity was strongly inhibited by the presence of ADP-hydroxymethylpyrrolidine diol, a specific inhibitor of PARG (*55*).

In Figure 9, material released from ADP-ribose polymers by anion exchange HPLC was analyzed. Extracts from a strain containing construct B were incubated with [^{32}P]ADP-ribose polymers (*56*), and a portion was analyzed by anion exchange HPLC as described. The elution times for AMP, ADPR, and PR-AMP are indicated by arrows. The material analyzed was PARG expressed in *E. coli*. The results indicated that the material released from ADP-ribose polymers

is exclusively ADP-ribose by strong anion exchange HPLC (Figure 9), demonstrating that the cell extracts did not contain any other ADP-ribose polymer-degrading enzymes such as phosphodiesterase, which catalyzes the formation of AMP and phosphoribosyl-AMP (*57*).

Anion exchange HPLC utilized a Whatman Partisil SAX column equilibrated with 7 mM potassium phosphate buffer pH 4.0, at a flow rate of 1 ml/min. The sample was diluted in the same buffer, applied to the column, and eluted with a 30-min linear gradient from 7 mM potassium phosphate buffer, pH 4.0 to 250 mM potassium phosphate buffer, 0.5 M KCl, pH 4.0.

To determine the size of the expressed enzymatic activity, an activity gel assay (*58*) was used. Activity gel assays for bPARG were done by casting polyacrylamide gels with automodified PARP containing [^{32}P]ADP-ribose polymers as described previously (*59*). Following electrophoresis, PARG was renatured by incubating the gels at 25°C in 5 volumes of 50 mM sodium phosphate buffer, pH 7.5, 50 mM NaCl, 10% glycerol, 1% Triton X-100, 10 mM β-mercaptoethanol, changing the buffer every 3 h for a total of five changes. After an additional incubation at 37°C for 3 h, gels were dried, and PARG activity was detected following autoradiography as a clear band on a black background. Cell extracts containing PARG fused to GST were examined for binding to glutathione-Sepharose 4B (GSH-Sepharose) (Pharmacia Biotech Inc.) according to the specifications of the manufacturer. No bands were produced from extracts from the IPTG-induced pTrcHisB vector that did not contain an insert. Extracts from cells transformed with a construct containing a PARG insert showed bands at approximately 115 and 59 kDa (Figure 8). During storage at 4°C, cell extracts lost activity migrating at the higher molecular weight, while the activity at approximately 59 kDa increased.

Expression of bPARG in the pTrcHisB expression vector did not result in detectable amounts of protein by staining the Coomassie Blue. Thus, another construction was designed to overexpress a 69-kDa carboxyl-terminal region of the PARG as a fusion with GST, which allows convenient protein purification by affinity chromatography on a GSH-Sepharose column. Two hours after induction with IPTG, strong expression of a protein migrating at approximately 90 kDa was observed. This protein bound to GSH-Sepharose and was eluted by GSH. The construct contained a thrombin cleavage site between the GST and the 69-kDa region of PARG, and the treatment of the material bound to GSH-Sepharose with thrombin resulted in the release of a protein that migrated at approximately 59 kDa. This result suggests that the protein purified from the bovine thymus may be larger than suggested by its migration on SDS-PAGE. The

result of this experiment is presented in Figure 10. Lane 1 shows extract from uninduced cells; lane 2 shows extract from cells induced with 1 mM IPTG for 2 hours; lane 3 shows proteins in extracts from cells shown in lane 2 that bound to GSH-Sepharose; lane 4 shows material released from GSH-Sepharose by treatment with thrombin.

In addition to the GST fusion construct described above, several other GST fusion proteins have been made. Figure 11 shows the portions of the bovine PARG gene that have been expressed. The top line represents the structure of bovine PARG mRNA containing the open reading frame encoding the 111 kDa PARG protein. The different parts of PARG that have been cloned in expression vectors are represented with the size of the resulting expressed recombinant proteins. The expression of the 65 kDa catalytic domain of PARG (starting at the amino acid MNDV) in pGEX-2T as a fusion protein with glutathione-S-transferase (29 kDa) is detailed. Among the constructs, only the clone designed to express a protein of 69 kDa starting at amino acid +380 from the sequence of bovine PARG ($bPARG_{MNDV}$) allowed high level expression as a fusion protein with glutathione-S transferase (GST). A 1.8 kb PCR *Eco*RI fragment encoding for the 65 kDa catalytic domain of PARG was cloned into the *Eco*RI site of pGEX-2T giving pGEX-2T-$bPARG_{MNDV}$. This construction results in the expression of a fused polypeptide consisting of the sequence of GST. Amino acids derived from the polylinker and thrombin site and the 65 kDa domain (Figure 12).

In addition to various constructs designed to express PARG in *E. coli*, a recombinant baculovirus expressing a functional PARG has been constructed using the methodology of Summers and Smith as set out in U.S. Patent 4,879,236 which is specifically incorporated herein by reference.

$bPARG_{MNDV}$ was cloned in baculovirus transfer vector pVL1393 using the EcoRI site. The recombinant vector was constructed as follows. An insert containing the cDNA sequence from position 1138-2946 of bovine PARG was prepared by subcloning a 1.8kb EcoRI fragment amplified from pWL30 using oligonucleotides CCAATTTGAAGGAGGAATTCCCGCCGCCACCATGAATGATGTGAATGCCAAACG ACCTGGA (SEQ ID NO: 34) and GCGTCTAGAATTCACTTGGCTCCTCAGGC (SEQ ID NO: 31, WIN15). The resulting fragment was inserted into the EcoRI site of the pVL1393 baculovirus transfer vector. The amplification introduced a Kozak consensus sequence (gaattcccgccgccaccATGAA SEQ ID NO: 35) at the start site of translation to enhance expression

of the recombinant protein. The resulting recombinant plasmid was cotransfected with linearize Baculogold™ baculovirus DNA (Pharmingen, San Diego, CA) into SF9 cells according to the manufacturers instructions. Recombinant viruses isolated using standard techniques. Overexpression of the recombinant protein was confirmed by Western blot and the results displayed in Figure 13 demonstrate that the 65 kDa domain expressed in *E. coli* contained enzymatic activity (lane 2) migrating with the same apparent molecular weight as the enzyme purified from bovine thymus (lane 1). Likewise, a construct expressing $bPARG_{MNDV}$ domain in SF9 insect cells infected with recombinant baculovirus showed activity (lane 4) migrating with the same apparent molecular weight.

Example 6 Northern Blot Analysis.

An surprising feature of the consensus full-length cDNA clone was that it predicted expression of a protein of approximately 111 kDa (Figure 3, SEQ ID NO: 1, and SEQ ID NO: 2), while the enzymatically active PARG from thymus had a molecular weight of approximately 59 kDa (Figure 2). To determine the size of the RNA transcript for PARG, total RNA and poly(A)+ RNA were isolated from bovine kidney (MDBK) cells and annealed using Clone 4 as the hybridization probe.

Total cytoplasmic RNA and poly(A)+ RNA were isolated from bovine kidney MDBK cells (ATCC #CCL22) using TRIzol reagent (Gibco/BRL) following the manufacturer's recommendations. After the RNA was fractionated, it was then transferred to nylon membranes and hybridized with Clone 4 (Figure 3) radiolabeled by a random hexamer priming method (21). The results are presented in Figure 4. Total RNA (5 μg, lanes 1A and 1B) and poly(A)+ RNA (4 μg, lanes 2A and 2B) were separated on a denaturing agarose gel (*60*). Panel A shows the ethidium bromide stained gel and panel B shows the autoradiogram of a Northern blot analysis using a random primed, ^{32}P-labeled DNA probe constructed from Clone 4 (Figure 3). A single transcript of approximately 4.3 kb was detected in the poly(A)+ RNA (Figure 4, lane 2). Thus, the transcript size was consistent with the expression of a 111 kDa PARG protein.

Example 7 Southern Blot Analysis of PARG Genomic Complexity.

Previous studies have reported that PARG isolated from nuclear fractions had a molecular weight of approximately 75 kDa (*61*), while PARG isolated from whole cell homogenates or

postnuclear supernatant fractions had a molecular weight of approximately 59 kDa (*62*). These results suggest that either two or more genes may code for PARG or that proteolysis generates lower molecular weight forms from higher molecular weight forms. The cDNA isolated encoded a protein considerably larger than any PARG proteins previously described, consistent with the possibility that the different forms of PARG are derived from a single form by proteolytic cleavage. To test the hypothesis that PARG is encoded by a single copy gene, the genomic complexity of the PARG gene was analyzed by a Southern hybridization experiment.

Total genomic DNA was prepared from bovine thymus tissue as described previously (*63*) and DNA (10 μg) was digested with EcoRI, BglII, XbaI or PstI, fractionated on a 1% agarose gel, transferred to a nylon membrane (Hybond N+, Amersham), and hybridized using an 828 bp HindIII fragment of Clone 1 radiolabeled as described for clone 4 above (*64*). Pre-hybridizations and hybridizations were carried out at 42 °C in 50% formamide, 0.25 M sodium phosphate buffer, pH 7.2, 0.25 M NaCl, 7% SDS, 1 mM EDTA. The blot was annealed with a ^{32}P-labeled DNA probe corresponding to the carboxyl terminal region of the PARG protein.

The results of the Southern blot analysis are presented in Figure 7. Genomic DNA was digested with four different restriction enzymes, EcoRI (lane 1), BglII (lane 2), XbaI (lane 3) and PstI (lane 4), none of which cleave within the carboxyl terminal region of the PARG cDNA. Following electrophoresis, the restriction digests were subjected to hybridization with a probe that corresponded to the carboxyl terminal region of the PARG cDNA. The analysis displayed in Figure 7 shows that, in each restriction digest, the probe hybridized primarily with a single restriction fragment. The fainter signals likely reflect the presence of introns in the PARG gene. This result indicates that PARG is encoded by a single copy gene in the bovine genome.

Example 8 Isolation and Characterization of PARGs from Other Species.

The isolation and characterization of bovine cDNA encoding poly(ADP-ribose) glycohydrolase (PARG) has been described above. Using the information provided by the sequencing of bovine PARG, various tools were used, including public sequence databases searches and screening of cDNA libraries using PARG specific probes, to clone and sequence the cDNA and determine the primary structure of PARG from human, mouse, *Drosophila* and *Caernorhabditis elegans*. Mammalian sequences newly obtained using this combined strategy

show high sequence similarity to bovine PARG (bPARG), whereas the sequences of *Drosophila* and *C. elegans* only display significant homologies in the region responsible of the catalytic activity of the protein.

The strategy followed to obtain cDNAs coding for proteins with sequence similarity to bovine PARG is summarized in Figure 14. dBEST, GenBank, SwissProt and PIR databases were searched for PARG like sequences at the nucleotide or amino acid level using the programs BLASTn, TBLASTn (Altschul *et al.*, 1990) respectively, available at the NIH site on the Worldwide Web, and also included in the sequence analysis package from the Genetic Computer Group, Inc. (GCG) (Madison, WI), version 9.1. Both programs perform pair-wise sequence comparisons on multiple nucleotide or amino acid sequences. PARG multiple sequence comparisons obtained with these programs are very similar. Box-shading of the amino acids in the multi-sequence alignment was obtained using the program BOXSHADE (K. Hofmann and M.D. Baron). The first step involved extensive searching for sequences with bPARG similarity in various databases. As a result of this search several partial nucleotide sequences sharing extensive homologies with bPARG cDNA were obtained from the dBEST database (*65*). These sequences were the result of random cloning and sequencing of partial cDNAs clones obtained from mRNAs expressed in various tissues and organisms. Among them, partial cDNAs coding for PARG from human and mouse were available. One of these human clones was particularly interesting as its sequence (2500 bp long) overlapped the coding sequence of bovine PARG from aa470 to aa977 (Carboxy terminus end) and contained all the 3' untranslated region of the human PARG cDNA. This clone (No. 50859; GenBank accession number: H17209) was requested and freely obtained from the IMAGE Consortium (in collaboration with Washington University School of Medicine in St. Louis, MO and Merck & Co., info@image.llnl.gov). The sequence of the clone was then completed. This partial cDNA permitted design of a radiolabeled probe (fragment *Hind*III - *Kpn*I of 677 bp) specific to human PARG (SEQ ID NO: 36).

Example 9 Cloning and sequencing.

The cloning procedures used in this work generally known and are also described in details in the book, Molecular Cloning: A Laboratory Manual (Maniatis *et al.*, 1982). DNA sequencing was performed using the dideoxynucleotide method of Sanger (Sanger *et al.*, 1977). Chemical reagents were purchased from Sigma (St. Louis, MO). Restriction enzymes, T4 DNA

ligase were from New England Biolabs, Inc. (Beverly, MA), T7 DNA polymerase Sequenase from US Biochemical (Cleveland, OH), Calf Intestine Phosphatase from Boehringer, Mannheim (Indianapolis, IN). The phagemid pTZ18/19R is from Pharmacia (Piscataway, NJ). The labeled nucleotides α-[^{35}S]-dATP and α-[^{32}P]-dCTP were purchased from ICN (Costa Mesa, CA). Human thymus and murine liver 5'-stretch cDNA libraries cloned in the vector λgt 10 were from Clontech (Palo Alto, CA).

A single, isolated colony of *C600Hfl E. coli* strain was picked and grown in 5 ml of Luria-Bertani medium (LB) + 10 mM $MgSO_4$ + 0.2% maltose overnight at 37°C in a shaker. The bovine library lysate was diluted 1:250,000 and incubated with the *C600Hfl* bacterial overnight culture and 1X lambda dilution buffer. Next, LB soft top agar + 10 mM $MgSO_4$ was added, and the entire mixture was quickly poured onto 90 mm LB agar + 10 mM $MgSO_4$ plates. The plates were cooled briefly at room temperature to allow the inoculum to soak into the agar before they were incubated at 37°C for 6-7 hr. The number of clear plaques was counted to determine the titer.

Plaques containing the entire library that had been plated were transferred to nitrocellulose or nylon membranes. The filters were then washed in a 1.5 M NaCl/0.5 M NaOH solution to lyse the cells. This was followed by a 5 min wash in neutralizing solution (1.5 M NaCl/1 M Tris buffer pH 8). Finally, the filters were rinsed in 0.2X SSPE (30 mM NaCl/2 mM sodium phosphate buffer pH 7.2/0.2 mM EDTA) (Sambrook *et al.*, 1992). The filters were then dried and baked in a 80°C oven for 2 hr to fix the lysed plaques onto the filters.

Radioactive probes were prepared using a random hexamer priming method. Pre-hybridizations and hybridizations were carried out at 42°C in 50% formamide, 0.25 M sodium phosphate buffer, pH 7.2, 0.25 M NaCl, 7% SDS, 1 mM EDTA.

Example 10 Specific Methods used for Library Screening.

All the cloning procedures used in obtaining the additional PARG cDNAs and determining their sequences were performed essentially as described for the bovine PARG cDNA and sequence. Human thymus and murine liver 5'-stretch cDNA libraries cloned in the vector λgt 10 were from Clontech (Palo Alto, CA).

Library plating and titering: A single, isolated colony of C600Hfl *E. coli* strain was picked and grown in 5 ml of Luria-Bertani medium (LB) + 10 mM MgSO4 + 0.2% maltose

overnight at 37 °C in a shaker. The library lysate was diluted 1:250000 and incubated with the C600Nfl bacterial overnight culture and 1X lambda dilution buffer. Next, LB soft top agar + 10 mM MgSO4 was added, and the entire mixture was quickly poured onto 90 mm LB agar + 10 mM MgSO4 plates. The plates were cooled briefly at room temperature to allow the inoculum to soak into the agar before they were incubated at 37 °C for 6-7 hr. The number of clear plaques was counted to determine the titer.

Plaque lifts: Plaques containing the entire library that have been plated are transferred to nitrocellulose or nylon membranes. The filters are then washed in a 1.5 M NaCl/0.5 M NaOH solution to lyse the cells. This is followed by a 5 min wash in neutralizing solution (1.5 M NaCl/1 M Tris buffer pH 8). Finally, the filters are rinsed in 0.2X SSPE (30 mM NaCl/2 mM sodium phosphate buffer pH 7.2/0.2 mM EDTA) (*66*). The filters are then dried and baked in a 80 °C oven for 2 hr to fix the lysed plaques onto the filters.

Making a radioactive probe and Hybridizations: Radioactive probes were prepared using a random hexamer priming method. Pre-hybridizations and hybridizations were carried out at 42 °C in 50% formamide, 0.25 M sodium phosphate buffer, pH 7.2, 0.25 M NaCl, 7% SDS, 1 mM EDTA. This partial cDNA allowed to design a radiolabeled probe (fragment HindIII - KpnI of 750 bp long) specific to human PARG.

Example 11 Screening of a human thymus 5'-stretch cDNA libraries.

Multiple screenings of a human thymus 5'-stretch cDNA library were performed to complete the cloning of human PARG cDNA. For each screening a new probe was designed and used to screen approximately one million recombinants of the library. During each round of screening, overlapping clones were isolated at high stringency conditions and subcloned into the EcoRI site of pTZ18/19R phagemid using standard techniques. The different positive clones (J5, C, E1, E2, M, M', M", P', P", Of, 02) were characterized by restriction analysis, subcloned into the appropriate restriction sites of pTZ18/19R as necessary and sequenced in both strand using the dideoxynucleotide method. The probe used to complete the cloning of the human cDNA library is shown is SEQ ID NO: 37. Finally, a full-length cDNA sequence was assembled which encodes the human PARG. The sequence of the cDNA encoding human PARG is presented in the sequence listing as SEQ ID NO: 3 and the amino acid sequence of human PARG is presented in the sequence listing as SEQ ID NO: 4.

The human PARG sequence shares extensive amino acid sequence homologies with bovine PARG with more than 89% identity. The sequence similarity is also high at the nucleotide level particularly in the region coding for the protein (174ATG - TGA3104). Surprisingly the 5'-untranslated region of the human sequence displays a completely different sequence with an extensive sequence similarity with highly repeated polymorphic DNA sequences found in the human genome such as Alu repetitive elements or variable number of tandem repeats (VNTR).

Example 12 Screening of Mouse Liver 5' -stretch cDNA Libraries.

To isolate a PARG cDNA from the mouse liver cDNA library, a probe was designed from the human cDNA clone coding for PARG. Analysis of the bovine and human sequences revealed that PARG was highly conserved between these two species, suggesting that it might also be conserved in the mouse. Based on the restriction map of the human cDNA clone, a region in the human clone was selected, located where the active site of the protein is encoded, that exhibited near identity to its counterpart in the bovine clone. This region, consisting of approximately 800 bases, was excised from the entire human clone by digestion with the restriction endonuclease, *HindIII*, then purified by agarose gel separation and radiolabeled by random priming.

This probe was used to screen a mouse liver 5'-stretch cDNA library. One clone consistently hybridized with the probe. After two rounds of screening to ensure the purity of the clone the 2.5 kb insert was subcloned into the plasmid pTZ19R and sequenced. Comparison with the sequence of bovine and human PARG showed that this clone had the partial sequence that has extensive similarities to the two other mammalian sequences covering almost entirely the coding region from nucleotide - 10 to a few nucleotides from the end of the coding region. A second screen was performed to obtain the missing part of the cDNA using a radioactive probe specifically designed to hybridize with the region the most 3' of the previous clone to increase the chance to get the missing part of the cDNA.

With this new probe, the same mouse liver cDNA library was screened to obtain a second clone, containing an insert that was about 3 kb. This clone was purified, subcloned and sequenced. The sequence showed that this second clone starts at amino acids 634, extends toward the stop codon to approximately 900 nucleotides into the 3' non-coding region.

A search of the dBEST database turned up one significant match to a 400 bp fragment cloned from mouse muscularis. This fragment had an exact match to the very tail end of the second clone and exceeded it by 34 bases. This extra extension contained the oligo A sequence as well as the polyadenylation signal. Because there was an exact match, the cDNA sequence was completed using this information coming from the database. The complete cDNA sequence of murine PARG is presented in the Sequence Listing as SEQ ID NO: 5 and SEQ ID NO: 6.

Example 13 Obtaining the Drosophila PARG cDNA.

Among the clones obtained from DNA databases searches were several clones from the *Drosophila* genome sequencing project (European Drosophila Genome Sequencing Consortium) as well as the *Drosophila* expression sequence TAG sequencing project (*67*). The EST clone was requested from the University of California Berkeley and obtained. Because the sequence published in the dBEST database was only partial, its sequence was completed in our laboratory and compared to a genomic sequence, part of the distal X chromosome of *Drosophila melanogaster* submitted by Murphy *et al.*, August 1997 which presumably contains the gene of *Drosophila* PARG. The 768 aa shares less homologies with only 40% identity (48% similarity) mainly located in the catalytic domain of the protein. The domain organization of the protein is also very different with an unknown domain of 20 kDa located Carboxy terminus of the highly conserved active domain. (See Figure 15). The sequence of the cDNA encoding the Drosophila PARG is presented in the Sequence Listing as SEQ ID NO: 7 and the amino acid sequence of the Drosophila PARG is presented in the Sequence Listing as SEQ ID NO: 8.

Example 14 Obtaining the C. elegans PARG Sequence.

This sequence has been obtained by searching the GenBank database with the mammalian PARG protein sequence. A sequence with PARG similarity was found in the cosmid F20C5 (Accession number: Z68161, SEQ ID NO: 38) derived from the *C. elegans* genomic DNA (*68*). The overall sequence conservation (726aa, MW 83129 Da) with the other PARG sequences is as follows: 32% similarity and 22% identity with the mammalian PARG and 39% similarity and 30% identity with the *Drosophila* PARG. The sequence is presented in the Sequence Listing as SEQ ID NO: 38 (Genbank accession number CEF20C5). SEQ ID NO: 38 contains 12 exons as follows: exon 1 from 3591 to 3635; exon 2 from 3681 to 4121; exon 3 from 5065 to 5235; exon

4 from 5930 to 6152; exon 5 from 6200 to 6267; exon 6 from 7246 to 7338; exon 7 from 7386 to 7553; exon 8 from 7738 to 7853; exon 9 from 8153 to 8435; exon 10 from 8487 to: 8610; exon 11 from 8662 to 8952; and exon 12 from 9383 to: 9540. The coding sequence of the CePARG protein, which is publicly available from Accession number: Z68161, is referred to in the Sequence Listing as SEQ ID NO: 9. Its corresponding amino acid sequence is referred to in the Sequence Listing as SEQ ID NO: 10. The amino acid sequence of the C. elegans PARG is presented on the alignment (Figure 16)

Example 15 Cloning and Overproduction of the Carboxyl-terminus 69 kDa Domain of Bovine PARG (bPARG) in E. coli.

As described, above, bovine PARG is encoded by a messenger of 4 kb predicting a protein of 110 kDa, almost twice the size of the purified enzyme (65 kDa). It is also demonstrated that bPARG can be expressed in E. coli as an active enzyme either as a 110 kDa or a 65 kDa protein. This result combined with other evidence implies that the active site of PARG is located in the carboxyl-terminal part of the protein. Figure 11 is a schematic representation of the different clones we have expressed in bacteria. Among them, only the clone designed to express a protein of 69 kDa starting at the amino acid +380 from the sequence of bovine PARG ($bPARG_{MNDV}$) allowed high level expression as a fusion protein with glutathione-S transferase (GST).

The heterologous expression of $bPARG_{MNDV}$ was conducted as represented in Figure 12. The 1.8 kb cDNA encoding the 69 kDa carboxyl-terminal part of bovine PARG was amplified by PCR and cloned in the EcoRI site of pGEX-2T vector (Pharmacia) in fusion with GST giving the pGEX-2T-$bPARG_{MNDV}$ plasmid. *E. coli* NM522 cells transformed with the pGEX-2T-$bPARG_{MNDV}$ were induced by addition of IPTG, resulting in expression of a 90 kDa fusion protein. The fusion protein can be conveniently purified using Glutathione-Sepharose and the $bPARG_{MNDV}$ can be released by treatment with thrombin while the GST protein remains bound to the beads of GSH-Sepharose. In this manner milligram amounts of protein can be routinely obtained.

Example 16 Characterization of the purified 65 kDa Domain and the Generation of Antibodies.

The purified $bPARG_{MNDV}$ was characterized by activity gel assays (*69*) by casting polyacrylamide gels with automodified PARP containing [^{32}P]ADP-ribose polymers. The results demonstrate that the 65 kDa domain expressed in *E. coli* contained enzymatic activity migrating with the same apparent molecular weight as the enzyme purified from bovine thymus. Likewise, a construction expressing $bPARG_{MNDV}$ domain in SF9 insect cells infected with recombinant baculovirus showed activity migrating with the same apparent molecular weight.

The availability of PARG cDNA allows the development of new molecular tools to study this enzyme in its cellular context. Until this work, it was not possible to obtain PARG in sufficient quantities to produce antibodies against the protein. The antibody raised against bovine PARG is able to recognize PARG from other organisms and, thus, will be valuable in characterizing PARG *in vivo* under defined physiological conditions in many different organisms.

Antibodies against $bPARG_{MNDV}$ overexpressed in *E. coli* were raised in rabbits using the procedure described by Vaitukaitis (*70*). Specific high affinity antibodies are generated by administration of small doses of immunogens intradermally over a wide anatomic area of the animal. Rabbits were immunized by three injections of 10-50 μg of the Mr 65,000 protein band excised from a preparative SDS polyacrylamide gel. Titer and affinity of sera harvested weekly were followed by conventional methods. Peak affinity was attained in 8 to 10 weeks after primary immunization. For each animal, a preimmune serum was retained as a control.

Figure 17 shows a Western blot experiment demonstrating the specificity of the resulting PARG anti-serum against the purified bPARG from thymus (lane 1), SF9 protein extract expressing 65 kDa-$bPARG_{MNDV}$ in recombinant baculovirus (lane 2), recombinant 65 kDa-$PARG_{MNDV}$ purified by treatment with thrombin from GSH-Sepharose (lane 3), and an *E. coli* crude extract expressing the fusion protein GST-65 kDa-$PARG_{MNDV}$ (lane 4). The pre-immune serum did not show reactivity against any of these fractions even at a low dilution (1/250).

Antibodies directed against the 45 kDa –terminal have also been generated using the same strategy used to generate antibodies against the catalytic domain. This involved the overexpression of the 45 kDa protein domain in *E. coli* in a construct designed for easy purification, followed by injection of the purified protein into rabbits. The heterologous

expression of PARG45 was conducted by cloning a part (1.1 kb) of the coding region of the cDNA, generated by PCR amplification of the region located between the ATG(267) codon and nucleotide 1400 in the bovine sequence, into the Eco RI site of the bacterial expression vector pGEX-2T (Pharmacia) in fusion with glutathione-S-transferase. *E. coli* NM522 cells transformed with this construct were induced by addition of IPTG, resulting in expression of a 72 kDa fusion protein. The fusion protein was purified using glutathione Sepharose and the PARG45 was released by treatment with thrombin, while the GST protein remained bound to the GSH Sepharose beads. In this manner milligram amounts of protein were obtained. Antibodies against PARG45 overexpressed in E. coli were raised in rabbits using the procedure described (*71*). Specific high affinity antibodies were generated by administration of small doses of immunogens subcutaneously over a wide area of the animal. Rabbits were immunized by three injections of 10-50 μg of the 45 kDa protein band excised from a preparative SDS-polyacrylamide gel. Titer and affinity of sera harvested weekly were followed by conventional methods. Peak affinity was attained in 8 to 10 weeks after primary immunization. For each animal, a preimmune serum was retained as a control.

Example 17 Conservation of PARG in Tissues and Organisms.

Tissue and cell extracts from different origins were homogenized in a cold hypotonic lysis buffer containing a cocktail of protease inhibitors and sonicated. SDS and β-mercaptoethanol were added to insure inactivation of any remaining active proteases. Thirty μg of protein from each extract was analyzed by Western-blot using the anti-PARG antibody (Figure 18). In all of the fractions from bovine tissues, PARG was observed as a major band at 65 kDa. However, less intense, discrete proteins of higher molecular weight were also detected. These proteins may correspond to different forms of PARG; the band of highest molecular weight (about 115 kDa) found in thymus extract likely corresponds to the full-length of PARG (111 kDa) as deduced from the cDNA. Multiple species were detected in cell extracts from mouse fibroblasts, rat PC 12 cells, and SF9 insect cells. This result shows that the sequence of PARG is well conserved phylogenetically. Moreover, the conservation includes multiple molecular forms of the protein.

Example 18 Regulation of the expression of PARG.

In the metabolism of ADP-ribose polymers, the activities of PARP and PARG are closely related. Soon after polymer has been synthesized by PARP following DNA damage, it is extensively degraded by PARG. The net result is that the polymer has a very short half life. The close relationship between the two proteins suggests a possible mode of regulation in which PARG expression depends on the presence of PARP. In order to test if the presence or the absence of PARP influences the expression of PARG, a Western Blot experiment was performed with cell extracts from mouse fibroblasts of different PARP genotypes (*72*).

Cell extract (30 μg) from mouse cells with PARP +/+, PARP +/- and PARP-/- genotypes were separated by SDS-PAGE, transferred to a membrane and probed with the antibodies indicated. The results are shown in Figure 19. In Figure 19, purified PARG (50 ng) from bovine thymus (lane 1), 30 μg of protein of a total extract from PARG recombinant baculovirus infected SF9 cells (lane 2), 150 ng of purified recombinant PARG produced in the bacteria (lane 3) and 30 μg of protein of a crude extract from *E. coli* NM522 transformed with pGEX-2T-bPARG$_{MNDV}$ 2 h after induction by IPTG (lane 4) were separated on a 0.1% SDS-12% polyacrylamide gel, then transferred on nitrocellulose, and incubated with a 1/5000 dilution of the rabbit polyclonal antiserum raised against the 65 kDa domain of bPARG. Proteins were revealed by immunofluorescence with the ECL detection kit (Amersham) and autoradiography. Panel A is a western blot of PARP in cells of varying PARP genotype showing the results of the analysis using anti-PARP antibodies. The amount of PARP expressed varies as expected dependant upon the genotype of the cell line with the PARP-/- cell line producing no detectable amount of PARP. Panel B is western blot of PARG from various tissues using an anti-PARG antibody. It shows that the level of PARP is variable. The amount of PARG present in the cell extracts was not dependent upon the PARG genotype of the cell. Further support for this view is provided by the results of the PARG activity assay presented in panel C. The specific activity of PARG detected in the extracts showed no significant difference among the three genotypes.

Example 19 Preparation PARG Gene Ablation (Knockout) Animals.

One embodiment of the present invention is experimental animals with targeted mutations in the PARG gene. These animals may be constructed using standard techniques and the cDNA sequence of the PARG. In the following example, a mouse containing a targeted mutation in the

PARG gene is constructed. Those skilled in the art will readily appreciate that other experimental animals, including but not limited to rats, guinea pigs, hamsters and the like, may be constructed using similar techniques. The construction of animals with disrupted genes may be accomplished using standard techniques such as those described by Moreadith (*73*). Further, cells lines, construction kits, and protocols for knockout mice are available from commercial suppliers such as Stratagene (Stratagene 1999 catalog, La Jolla, CA). Commercial services such as Lexicon Genetics (The Woodlands, TX) and Chrysalis DNX Transgenic Sciences (Princeton, NJ) also offer complete ES cell knockout mice production services.

A genomic clone of the murine PARG enzyme may be isolated from a genomic library by screening with a probe derived from the cDNA sequence of PARG. A mouse 129/SV genomic library (Stratagene) containing mouse genomic sequences in λ phage was screened using a 2.49 kb fragment of the mouse PARG cDNA as a probe. A partial restriction map of one positive clone thus isolated, R1, is provided in Figure 20. The R1 clone contains the genomic sequence corresponding to the 5′-most end of the murine cDNA. The clone was subcloned into pBluescript as three fragments. The plasmid containing the 5′-end contained a 2.8 kb NotI-EcoRI fragment and was designated p2.8R. The fragment containing the central portion contained a 3.5 kb EcoRI fragment and was designated p3.5R. The plasmid containing the 3′-end of the gene contained a 7,0 kb EcoRI-NotI fragment and was designated 7.0R. Sequencing the resulting plasmids revealed that p2.8R contained no sequences corresponding to the cDNA, p3.5R contains a 1.5 kb promoter region and/or untranslated region and exon 1 coding for 72 amino acids including the initiation ATG codon and p7.0R contains at least 4 additional exons. Gene targeting vectors may be constructed using both p3.5R and p7.0R.

A gene targeting vector may contain one or more selection genes flanked by genomic sequences. The targeting vector is introduced into the genome by homologous recombination resulting in the incorporation of the selection gene into the genome of the cell. The mouse PARG gene was targeted using a "conditional" inactivation procedure outlined in Figure 21. This approach allows the production of viable animals even if the disrupted gene results in a lethal phenotype since the gene is not disrupted until a second "conditional" recombination event is induced.

A lox-P sequence may be inserted into the first intron. A cassette expressing the neomycin resistance gene (*neo*) and the thymidine kinase gene (*TK*) flanked by two additional

lox-P sites may be placed in intron 2. In the presence of Cre recombinase, recombination will occur between two lox-P sites thereby deleting the genomic sequences present between the sites. A MC1-DTA cassette is ligated at the 3′-end of the vector to reduce random integration of the vector into the genome.

The targeting vector may be introduced into embryonic stem cells by any method known to those skilled in the art such as transfection, lipofection or electroporation. In a preferred embodiment, the targeting vector will be introduced into embryonic stem cells by electroporation. After homologous recombination, cells containing the *neo* gene will be selected for using G418. Selected cells will then be analyzed by PCR and Southern blot.

To generate mutant alleles of PARG, the positive embryonic stem cell clones identified will be transfected with a plasmid expressing Cre recombinase. The action of Cre recombinase will result in three different mutant alleles. Mutant allele I contains a deletion in exon 2 but still maintains the selection genes *neo* and *TK*. Mutant allele II contains the genomic sequence for exon 2 flanked by two lox-P sites (exon 2 is said to be "floxed") and does not contain the selection genes. Mutant allele III has a deletion of the genomic sequences and does not contain the selection genes.

Mice containing each of the three mutant alleles may be constructed by microinjecting embryonic stem cells containing the mutant allele into blastocytes resulting in the production of chimeric and mutant mice. Mice homozygous in mutant allele I or III will be null mutants in that they will be unable to express a functional PARG enzyme due to the loss of required genomic sequences. In the absence of Cre recombinase, mice containing mutant allele II will express a wild type protein. In the presence of the Cre recombinase, the PARG will lose exon 2, thus producing an inactive protein. To inactivate the gene, these mice will be bred to mice expressing Cre recombinase under the control of a tissue specific promoter. This will result in mice expressing PARG in some tissues and not expressing PARG in others. Mice homozygous in mutant allele II, will be valuable for evaluating the role of PARG in specific tissues.

Although the present invention has been described with reference to certain examples for purposes of clarification and illustration. It should be appreciated that certain obvious improvements and modifications can be practiced within the scope of the appended claims and their equivalents. Other embodiments and uses of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein.

All U.S. Patents GenBank sequence listings, and other references noted herein for whatever reason are specifically incorporated by reference. The specification and examples should be considered exemplary only with the true scope and spirit of the invention indicated by the following claims.

1. Miwa, M. et al. (1971) J. Biol. Chem. 246, 6362-6364; Ueda, K. *et al.* (1972) *Biochem. Biophys. Res. Commun.*46, 516-523

2. Oka, J. *et al.* (1984) *J. Biol. Chem.* 259, 986-995.

3. Gaal, J. *et al.* (1987) *Trends in Biol. Sci.*, 12, 129.

4. Nudka, N. *et al.* (1980) *Eur. J. Biochem.* 105, 525-530; Jacobson, E. *et al.* (1985) *Carcinogenesis* 6, 715-718; Küpper, J. *et al.* (1990) *J. Biol. Chem.* 265, 18721-18724; Ding, R. *et al.* (1992) *J. Biol. Chem.* 267, 12804-12812.

5. Jacobson, E. *et al.* (1985) in *ADP-ribosylation of Proteins* (Althaus, F.R., Hilz, H., and Shall, S., eds) pp. 277-283, Springer-Verlag, Berlin; Lubet, R. *et al.* (1986) *Carcinogenesis* 7, 71-75; Kasid, U. *et al.* (1986) *Carcinogenesis* 7, 327-330.

6. Berger, N. (1985) *Radiat. Res.* 101, 4-15.

7. Kaufmann, S. *et al.* (1993) *Cancer Res.* 53, 3976-3985; Lazebnik, Y. *et al.* (1994) *Nature* 371, 346-347.

8. de Murcia, G. *et al.* (1994) *Trends Biochem. Sci.* 19, 172-176.

9. BioWorld Today, Apr. 29, 1994, p. 3.

10. Thomassin, H. *et al.* (1990) *Nucleic Acids Res.* 18, 4691-4694.

11. Karlin, Samuel and Stephen F. Altschul (1993). Proc. Natl. Acad. Sci. USA 90:5873-7

12. Henikoff and Henikoff Proc. Natl. Acad. Sci. USA 89:10915-19, 1992

13. Altschul, Stephen F. (1991). J. Mol. Biol. 219:555-65.

14. Cohen, J. S., 1989; Weintraub, H. M., 1990

15. N. Sarver et al., 1990

16. Orlandi, R. et al., (1989) Proc. Natl. Acad. Sci. 86, 3833-3837; Winter, G et al., (1991) Nature 349, 293-299

17. R. Sobol and K. Scanlon eds., available at www.appleton-lange.com.

18. e.g., at least, but not limited to US5797870, US5804383, US5670161, US5645829, US5741486, US5836905, US5843069, US5827216, US5871464, US5702384, US5810888, US5787900, US5752515, US5674192, US5658955, US5656465, US5547932, US5873904,

US5792651, US5772888, US5641750, US5641749, and US5626561

19. Kaltenbock, B. *et al.* (1998) *Biotechniques*, 24, 202-206.

20. Moran, P. *et al.* (1998) *Biotechniques*, 24, 206-212.

21. Orita, M. *et al.* (1989) *PNAS USA*, 86, 2766-2770.

22. Liu, Q. *et al.* (1998) *Biotechniques*, 24, 140-147.

23. *Proceeding from the Sixth International Symposium on Human Identification 1995* (ISBN 1-882274-55-5).

24. Kohler and Milstein, *Nature*, 256:495-497 (1975).

25. Geysen, *et al.* (1984) *Proc. Natl. Acad. Sci. U S A* 81:3998-4002.

26. Ménard, *et al.* (1987) *Biochem.* Cell Biol. 65, 668-673.

27. Hatakeyama, K. *et al.* (1986) *J. Biol. Chem.* 261, 14902-14911.

28. Thomassin, H. *et al.* (1990) *Nucleic Acids Res.* 18, 4691-4694.

29. Ménard, L. *et al.* (1987) *Biochem. Cell Biol.* 65, 668-673.

30. Bradford, M. (1976) *Anal. Biochem.* 72, 248-254.

31. Laemmli, U. (1970) *Nature* 227, 680-685.

32. Althaus, F. *et al.* (1987) *Molecular Biology, Biochemistry and Biophysics*, Vol. 37, Springer-Verlag, Berlin.

33. Mead, D. *et al.* (1986) *Protein Eng.* 1, 67-74.

34. Shaw, *et al.* (1986) Cell 46, 659-667.

35. Kozak (1987) *Nucleic Acids Res.* 15, 8125-8148.

36. Wilson, *et al.* (1994) *Nature (London)* 368, 32-38.

37. Brendel, *et al.* (1992) *Proc. Natl. Acad. Sci. U S A* 89, 2002-2006.

38. Robbins, *et al.* (1991) *Cell* 64, 615-623.

39. Schreiber, *et al.* (1992) *EMBO J.* 11, 3263-3269.

40. Uchida, K. *et al.* (1987) *Biochem. Biophys. Res. Commun.* 148, 617-622; Cherney, B. *et al.* (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 8370-8374; Kurosaki, T. *et al.* (1987) *J. Biol. Chem.* 262, 15990-15997.

41. Huppi, K. *et al.* (1989) *Nucleic Acids Res.* 17, 3387-3401.

42. Saito, I. *et al.* (1990) *Gene (Amst.)* 90, 249-254.

43. Ittel, M.-E. *et al.* (1991) *Gene (Amst.)* 102, 157-164.

44. Saulier-Le Drean, B. (1992) *Poly(ADP-ribose) Polymerase in Xenopus lae-vis*, Ph.D. thesis, Université De Rennes, France.

45. Uchida, K. *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90, 3481-3485.

46. Masutani, M. *et al.* (1994) *Eur. J. Biochem.* 220, 607-614.

47. Schreiber, *et al.* (1992) *EMBO J.* 11, 3263-3269.

48. Lazebnik, *et al.* (1994) *Nature (London)* 371, 346-347.

49. Brochu, *et al.* (1994) *Biochem. Biophys. Acta* 1219, 342-350.

50. Tanuma, *et al.* (1986) *J. Biol. Chem.* 261, 965-969; Uchida, *et al.* (1993) *J. Biol. Chem.* 268, 3194-3200.

51. Maruta, *et al.* (1991) *Biochemistry* 30, 5907-5912.

52. Ménard, L. *et al.* (1987) *Biochem. Cell Biol.* 65, 668-673.

53. Ménard, L. *et al.* (1987) *Biochem. Cell Biol.* 65, 668-673.

54. Slama, J. *et al.* (1995) *J. Med. Chem.* 38, 389-393; Slama, J. *et al.* (1995) *J. Med. Chem.* 38, 4332-4336.

55. Slama, J. *et al.* (1995) *J. Med. Chem.* 38, 389-393; Slama, J. *et al.* (1995) *J. Med. Chem.* 38, 4332-4336.

56. Ménard, L. *et al.* (1987) *Biochem. Cell Biol.* 65, 668-673.

57. Althaus, F. *et al.* (1987) *Molecular Biology, Biochemistry and Biophysics*, Vol. 37, Springer-Verlag, Berlin.

58. Brochu, G. *et al.* (1994) *Anal. Biochem.* 218, 265-272.

59. Brochu, G. *et al.* (1994) *Anal. Biochem.* 218, 265-272.

60. Moreadith, *et al.* (1997) *J. Mol. Med.* 75, 208-216.

61. Tanuma, *et al.* (1986) *J. Biol. Chem.* 261, 965-969 and Uchida, *et al.* (1993) *J. Biol. Chem.* 268, 3194-3200.

62. Hatakeyama, *et al.*. (1986) *J. Biol. Chem.* 261, 14902-14911; Thomassin, *et al.* (1990) *Nucleic Acids Res.* 18, 4691-4694; and Maruta, *et al.* (1991) *Biochemistry* 30, 5907-5912.

63. Sambrook, *et al.* (1992) MOLECULAR CLONING: A LABORATORY MANUAL, Cold Spring Harbor Laboratory Cold Spring harbor, New York.

64. Feinberg, *et al.* (1983) *Anal. Biochem.* 132, 6-13.

65. Boguski, 1995

66. Sambrook, *et al.* (1992) MOLECULAR CLONING: A LABORATORY MANUAL, Cold Spring Harbor Laboratory Cold Spring Harbor, New York.

67. BDGP/HHMI Drosophila EST Project, University of California Berkeley, EST@fruitfly.berkeley.edu

68. Nematode Sequencing Project, Sanger Centre, Hinzton, Cambridge CB10 1RQ, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jes@sanger.ac.uk

69. Brochu G. *et al.* (1994) *Anal. Biochem.* 218, 265-272.

70. Vaitukaitis (1981) *Methods in Enzymology* 73, 46-52.

71. Vaitukaitis (1981) *Methods in Enzymology* 73, 46-52.

72. Wang, *et al.* (1995) *Genes & Dev.* 9, 509-520.

73. Moreadith, *et al.* (1997) *J. Mol. Med.* 75, 208-216.